

2016

Towards Improving The Evaluation Of Speech Production Deficits In Chronic Stroke

Artemis Alexandra Basilakos
University of South Carolina

Follow this and additional works at: <https://scholarcommons.sc.edu/etd>



Part of the [Communication Sciences and Disorders Commons](#)

Recommended Citation

Basilakos, A. A.(2016). *Towards Improving The Evaluation Of Speech Production Deficits In Chronic Stroke*. (Doctoral dissertation). Retrieved from <https://scholarcommons.sc.edu/etd/3771>

This Open Access Dissertation is brought to you by Scholar Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact dillarda@mailbox.sc.edu.

TOWARDS IMPROVING THE EVALUATION OF SPEECH PRODUCTION DEFICITS IN
CHRONIC STROKE

by

Artemis Alexandra Basilakos

Bachelor of Arts
University of South Carolina, 2009

Master of Arts
University of Georgia, 2012

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Communication Sciences and Disorders

Norman J. Arnold School of Public Health

University of South Carolina

2016

Accepted by:

Julius Fridriksson, Major Professor

Daniel Fogerty, Committee Member

Dirk den Ouden, Committee Member

Chris Rorden, Committee Member

Lacy Ford, Senior Vice Provost and Dean of Graduate Studies

© Copyright by Artemis Alexandra Basilakos, 2016
All Rights Reserved.

DEDICATION

To those who dedicated their time to participate in this study, I dedicate this to you.

ABSTRACT

One of the most devastating consequences of stroke is aphasia - a disorder that impairs communication across the domains of expressive and receptive language. In addition to language difficulties, stroke survivors may struggle with disruptions in speech motor planning and/or execution processes (i.e., a motor speech disorder, MSD). The clinical management of MSDs has been challenged by debates regarding their theoretical nature and clinical manifestations. This is especially true for differentiating speech production errors that can be attributed to aphasia (i.e., phonemic paraphasias) from lower-level motor planning/programming impairments (i.e., articulation errors that occur in apraxia of speech; AOS). Therefore, the purposes of this study were 1) to identify objective measures that have the greatest discriminative weight in diagnostic classification of AOS, and 2) using neuroimaging, to localize patterns of brain damage predictive of these behaviors.

Method: Stroke survivors ($N=58$; 21 female; mean age= 61.03 ± 10.01 ; months post-onset= 66.07 ± 52.93) were recruited as part of a larger study. Participants completed a thorough battery of speech and language testing and underwent a series of magnetic resonance imaging (MRI) sequences. Objective, acoustic measures were obtained from three connected speech samples. These variables quantified inter-articulatory planning, speech rhythm and prosody, and speech fluency. The number of phonemic and distortion errors per sample was also quantified. All measures were analyzed for group differences, and variables were subject to a linear discriminant analysis (LDA) to determine which

served as the best predictor of AOS. MRI data were analyzed with voxel-based lesion-symptom mapping and connectome-symptom mapping to relate patterns of cortical necrosis and white matter compromise to different aspects of disordered speech.

Results: Participants with both AOS and aphasia generally demonstrated significantly poorer performance across all production measures when compared to those with aphasia as their only impairment, and compared to those with no detectable speech or language impairment. The LDA model with the greatest classification accuracy correctly predicted 90.7% of cases. Neuroimaging analysis indicated that damage to mostly unique regions of the pre- and post-central gyri, the supramarginal gyrus, and white matter connections between these regions and subcortical structures was related to impaired speech production.

Conclusions: Results support and build upon recent studies that have sought to improve the assessment of post-stroke speech production. Findings are discussed with regard to contemporary models of speech production, guided by the overarching goal of refining the clinical evaluation and theoretical explanations of AOS.

TABLE OF CONTENTS

DEDICATION	iii
ABSTRACT.....	iv
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF ABBREVIATIONS.....	xii
CHAPTER 1: INTRODUCTION	1
1.1 LIMITATIONS OF CURRENT ASSESSMENTS	1
1.2 SPEECH PRODUCTION IMPAIRMENTS IN POST-STROKE INDIVIDUALS.....	4
1.3 NEUROANATOMICAL CORRELATES OF AOS AND APHASIA	10
1.4 MODELS OF SPEECH PRODUCTION	19
1.5 CHAPTER SUMMARY.....	25
1.6 STUDY PURPOSE.....	26
1.7 SPECIFIC AIMS AND HYPOTHESES	28
1.8 PILOT DATA	29
CHAPTER 2: METHOD	38
2.1 PARTICIPANTS	38
2.2 SPEECH AND LANGUAGE TESTING	41
2.3 NEUROIMAGING DATA.....	42
2.4 PROCEDURE.....	43

2.5 DATA ANALYSES.....	57
CHAPTER 3: RESULTS	64
3.1 BEHAVIORAL DATA: GROUP PERFORMANCE ACROSS MEASURES.....	64
3.2 SUMMARY OF BEHAVIORAL MEASURES	81
3.3 CLINICAL CLASSIFICATION USING LINEAR DISCRIMINANT ANALYSIS.....	82
3.4 NEUROIMAGING RESULTS: VLSM AND CONNECTOME ANALYSES ...	89
3.5 SUMMARY OF ALL RESULTS.....	101
CHAPTER 4: DISCUSSION.....	104
4.1 GENERAL DISCUSSION	104
4.2 SPECIFIC AIMS 1A & 1B: IMPROVING THE ACCURACY OF DIAGNOSTIC CLASSIFICATION.....	105
4.3 SPECIFIC AIM 1C: EXPLORING THE EXISTENCE OF AOS SUBTYPES	111
4.4 SPECIFIC AIM 2: BRAIN REGIONS AND CONNECTIONS THAT SUPPORT FLUENT SPEECH	112
4.5 THEORETICAL IMPLICATIONS	117
4.6 FINAL REMARKS	127
REFERENCES	129
APPENDIX A: THE APRAXIA OF SPEECH RATING SCALE	144
APPENDIX B: MOTOR SPEECH EVALUATION.....	145
APPENDIX C: SELECTED PNT ITEMS FOR WORD-LEVEL SCORING.....	146
APPENDIX D: LINGUAL DISTORTION ERRORS USED FOR NARROW TRANSCRIPTION	147
APPENDIX E: REGIONS CHOSEN FOR EXPLORATORY CONNECTOME ANALYSIS.....	148

APPENDIX F: MEANS (STANDARD DEVIATIONS) AND SIGNIFICANT GROUP DIFFERENCES FOR EACH VARIABLE	149
--	-----

APPENDIX G: CORRELATIONS BETWEEN ALL MEASURES	150
---	-----

LIST OF TABLES

Table 1.1 Speech production impairments in AOS and dysarthria	7
Table 1.2 Speech production impairments in AOS and aphasia.....	10
Table 1.3 Group classification results from the TwoStep cluster analysis	33
Table 1.4 Results from preliminary univariate VLSM analyses.....	34
Table 1.5 Significant inter-regional connections obtained from preliminary DTI analyses	35
Table 2.1 Characteristics of the study sample	40
Table 2.2 EMS variables used in analyses	54
Table 2.3 Measures used to rate phonemic, phonetic, and prosodic impairments.....	56
Table 3.1 Results from univariate ANOVAs for each frequency band	77
Table 3.2 Pairwise comparisons for band prominence values between 1-32 Hz.....	77
Table 3.3 Correlation coefficients for the EMS measures and speech production deficits.....	79
Table 3.4 Correlations coefficients between WAB-AQ scores and all study measures	82
Table 3.5 Correlations between ASRS items and the variables with the greatest weight in diagnostic classification.....	88
Table 3.6 Regions implicated in VLSM analyses.....	91
Table 3.7 Fiber counts derived from structural connectomes.....	98
Table 3.8 Results from significant whole brain DTI analyses.....	100
Table 3.9 Correlation coefficients for regional fiber counts and speech production errors	102

LIST OF FIGURES

Figure 1.1 Results from preliminary MDS analysis	31
Figure 1.2 Rankings of predictor importance from the TwoStep cluster analysis.....	32
Figure 1.3 Results from the preliminary univariate VLSM analyses	34
Figure 1.4 DTI results from the preliminary ASRS Items	35
Figure 2.1 Example spectrogram from a speaker with AOS	47
Figure 2.2 Example EMS plot for a speaker with AOS.....	52
Figure 3.1 nPVI-V coefficient values for each group.....	65
Figure 3.2 Mean nPVI-V coefficients for each AOS “subgroup”	66
Figure 3.3 Box plots for rPVI-WS and rPVI-SW	67
Figure 3.4 VOT variability measured from the connected speech samples	69
Figure 3.5 VOT variability measured from word-level production.....	70
Figure 3.6 Proportions of phonemic errors and distortion errors.....	72
Figure 3.7 Comparison of speech fluency measures (based on standard scores)	73
Figure 3.8 Amplitude modulation spectra for each group	74
Figure 3.9 Peak frequency and deviation for each group	75
Figure 3.10 Mean peak amplitude for each group	76
Figure 3.11 Band prominence for each frequency band within the 1-32 Hz range	78
Figure 3.12 Correlations between speech sound errors, ASRS AOS severity and band prominence.....	80
Figure 3.13 MDS plot for the AOS-Aphasia group versus those without AOS	85

Figure 3.14 Discriminative weights obtained from the two models used to classify participants with AOS from those without AOS	85
Figure 3.15 Discriminative weights obtained from the models used to classify the AOS-Aphasia participants from the Aphasia Only participants	86
Figure 3.16 Results from significant VLSM analyses	90
Figure 3.17 Significant regions predictive of 4Hz and 16Hz band prominence.....	93
Figure 3.18 Mean fiber count plots for the inter-regional white matter connections selected for further inspection.....	96
Figure 3.19 Inter-regional fiber count connections used for exploratory connectome analyses	99

LIST OF ABBREVIATIONS

ABA-2.....	<i>Apraxia Battery for Adults</i> , 2 nd Edition
AG.....	Angular Gyrus
AOS.....	Apraxia of Speech
ASRS.....	Apraxia of Speech Rating Scale
BA.....	Brodmann Area
Caud.....	Caudate
CR.....	Corona Radiata
CST.....	Corticospinal Tract
DIVA.....	Directions into the Velocities of Articulators
DTI.....	Diffusion Tensor Imaging
DWI.....	Diffusion Weighted Imaging
EMS.....	Envelope Modulation Spectra
FOG.....	Fronto-orbital Gyrus
GODIVA.....	Gradient Order DIVA
GP.....	Globus Pallidus
HSFC.....	Hierarchical State Feedback Control
IFGpo.....	Inferior Frontal Gyrus <i>Pars Opercularis</i>
IFGpt.....	Inferior Frontal Gyrus <i>Pars Triangularis</i>
Ins.....	Insula
ITG.....	Inferior Temporal Gyrus

L	Left
p.....	Posterior
PCT	Pontine Crossing Tract
PG	Parietal Gyrus
PrCG	Precentral Gyrus
PVI.....	Pairwise Variability Index
MDS.....	Multidimensional Scaling
MFG	Middle Frontal Gyrus
MOG	Middle Occipital Gyrus
MRI.....	Magnetic Resonance Imaging
MTG.....	Middle Temporal Gyrus
PET	Positron Emission Tomography
PPAOS	Primary Progressive Apraxia of Speech
PoCG.....	Post-central Gyrus
PWI	Perfusion Weighted Imaging
s.....	Superior
SFC	State Feedback Control
SFG	Superior Frontal Gyrus
SLF.....	Superior Longitudinal Fasciculus
SMG.....	Supramarginal Gyrus
SPG	Superior Parietal Gyrus
SPGI.....	Superior Precentral Gyrus of the Insula
(Area) Spt	Boundary of Parietal & Temporal Areas at Posterior Sylvian Fissure

STG Superior Temporal Gyrus

TR Thalamic Radiation

VOT Voice Onset Time

R Right

VLSM Voxel-Based Lesion-Symptom Mapping

WAB (-R) Western Aphasia Battery (-Revised)

WPM Words per Minute

CHAPTER 1

INTRODUCTION

In the United States, stroke is one of the most common forms of non-traumatic brain injury, and the second leading cause of adult disability (NINDS, 2014). It is estimated that over one million individuals are currently living with aphasia, a disorder of language caused most frequently by stroke-induced damage to the dominant (typically left) hemisphere. Considering that approximately 100,000 new cases of aphasia are diagnosed per year, and given that projected costs for post-stroke care are expected to increase from \$71.55 billion to \$184.13 billion by 2030 (Rubin, 2014), management of post-stroke deficits is an important public health concern. The increase in costs associated with post-stroke care can be attributed to the growth in the aging population at risk for stroke, as well as the prevalence of stroke survivors who may be unable to return to work due to post-stroke deficits, including chronic speech and language impairments.

1.1 LIMITATIONS OF CURRENT ASSESSMENTS

Although several meta-analyses have shown that aphasia treatment can be successful (Brady, Kelly, Godwin, & Enderby, 2014; Robey, 1994, 1998), and many large-scale aphasia treatment studies continue today, evidence supporting treatments for post-stroke speech production impairments lags far behind (Ballard et al., 2015; Wambaugh, Duffy, McNeil, Robin, & Rogers, 2006). Management of post-stroke speech and language disorders hinges on an accurate diagnosis and description of the behavioral

impairment so treatments that target the underlying deficit can be implemented. Currently available psychometric assessments (e.g., the *Western Aphasia Battery-Revised, WAB-R*, Kertesz, 2007; the *Aphasia Diagnostic Profiles*, Helm-Estabrooks, 1992; and the *Boston Diagnostic Aphasia Examination*, Goodglass, Kaplan, & Barresi, 2000) are instrumental in the classification of aphasia type. However, these assessments do not provide a full picture of other factors that can affect speech production. For example, in addition to aphasia, patients may present with a motor speech disorder (MSD), i.e., apraxia of speech (AOS) or dysarthria. AOS and dysarthria are not caused by impairments in language processes; rather, they are impairments in planning and programming speech motor movements (AOS) and the control and execution of these movements (dysarthria). MSDs commonly co-occur with aphasia, and perceptually, these disorders share similar behavioral characteristics. Debates surrounding the behavioral manifestations of MSDs have impeded the adoption of uniform diagnostic criteria for these disorders, especially AOS. Consequently, there is a pressing need to improve diagnostic tools used to assess post-stroke patients, as doing so could ultimately inform future treatment studies regarding the nature of apraxic impairment, and how it should be treated to improve overall speech production (Ballard et al., 2016; Ballard et al., 2015; Haley, Jacks, & Cunningham, 2013; Haley, Jacks, de Riesthal, Abou-Khalil, & Roth, 2012; Strand, Duffy, Clark, & Josephs, 2014; Wambaugh et al., 2006; Wambaugh, Wright, Nessler, & Mauszycki, 2014). Aside from facilitating its clinical management, improving the assessment of AOS, and subsequently studying how brain damage impacts speech motor planning and programming can provide important information regarding the neuroanatomical structures and network connections that support speech production.

Assessments for AOS and dysarthria exist (Dabul, 2000; Enderby, 1980; Yorkston & Beukelman, 1981), but they are outdated with respect to current research, suffer from low reliability and validity, and are based on subjective judgment of speech samples. For example, the *Apraxia Battery for Adults-2nd Edition* (ABA-2; Dabul, 2000) often falsely classifies individuals with aphasia as having AOS because it does not provide clinicians with a guide to distinguish errors that can be attributed to aphasia (i.e., phonemic errors) from articulatory errors that are characteristic of AOS (McNeil, Pratt, & Fossett, 2004). Similarly, the *Frenchay Dysarthria Assessment* (Enderby, 1980) does not provide a means to differentiate dysarthric production impairments from those that can occur in AOS, as tasks (e.g., counting, sentence and word repetition) are simply rated on a severity scale of movement accuracy and intelligibility, and both AOS and dysarthria can impair those domains (Moser, Basilakos, Fillmore, & Fridriksson, 2016). To address these concerns, many studies have sought to identify behaviors and patterns of brain damage that reliably distinguish patients with aphasia from those that have both aphasia and an MSD, with the goal of establishing more objective methods for differential diagnosis (Basilakos, Rorden, Bonilha, Moser, & Fridriksson, 2015; Galluzzi, Bureca, Guariglia, & Romani, 2015; Graff-Radford et al., 2014; Haley et al., 2013; Haley et al., 2012; Haley & Martin, 2011; Itoh & Sasanuma, 1984; Johns & Darley, 1970; Kent & Rosenbek, 1983; Odell, McNeil, Rosenbek, & Hunter, 1991; Rosenbek, Wertz, & Darley, 1973; Seddoh et al., 1996; Strand et al., 2014; Vergis et al., 2014). More recently, an emphasis has been placed on establishing acoustic measures that can capture differences between phonemic errors (attributed to aphasia) and phonetic errors (attributed to an MSD; Ballard et al., 2016; Cunningham, Haley, & Jacks, 2015; Haley et al., 2012; Vergis

et al., 2014). In this manner, the use of acoustic measures reduces subjectivity in the description of speech production errors.

Taken together, research regarding the successful management of post-stroke communication deficits must consider the importance of diagnostic accuracy. Unfortunately this has remained a challenge, as there are currently no agreed upon methods for the differential diagnosis between aphasia and MSDs, and in particular, AOS. In turn, the literature is rife with studies that have adopted different definitions of AOS, as well as the criteria used to diagnose it. This is troublesome for clinical practice, as the lack of agreed upon theoretical and neuroanatomical bases of AOS has been an impediment to developing theoretically based and empirically supported treatments for this disorder.

1.2 SPEECH PRODUCTION IMPAIRMENTS IN POST-STROKE INDIVIDUALS

MSDs are impairments in articulatory motor planning and/or programming (AOS), or the motor execution and control (dysarthria) of articulation (Duffy, 2005). These disorders arise from impairments in processes that occur following lexical selection and phonological encoding, meaning they are not thought to reflect compromised linguistic processes (Haley et al., 2012; Itoh & Sasanuma, 1984; Jordan & Hillis, 2006; Seddoh et al., 1996; van der Merwe, 1997; Ziegler, Aichert, & Staiger, 2012). Specifically, AOS occurs due to impairments in the processes that are initiated following lexical selection and phonological encoding, at the stage of phonetic encoding. At this stage, motor codes and muscle commands are formulated from stored sensorimotor programs so that articulatory movements can be handled by the motor system for subsequent speech execution (van der Merwe, 1997). Although other

impairments may co-occur, the speech errors that characterize AOS cannot be attributed to deficient linguistic processes (i.e., aphasia) or impairments affecting the speech musculature (i.e., paralysis, paresis, spasticity, or discoordination that can occur in dysarthrias).

AOS is characterized by reduced speech rate, off-target and distorted articulation, false starts and restarts, and visible/audible groping of the articulators (McNeil, Robin, & Schmidt, 1997; Ogar, Slama, Dronkers, Amici, & Gorno-Tempini, 2005; Rosenbek, Kent, & Lapointe, 1984; Rosenbek, McNeil, & Aronson, 1984; Strand et al., 2014; van der Merwe, 1997). Suprasegmental aspects of production may also be affected, with prosodic impairments realized as increased intersegmental durations between phonemes, syllables, and words, as well as improper stress assignment to multisyllabic words or to words in sentences (Ballard et al., 2016; McNeil, Doyle, & Wambaugh, 2000; McNeil et al., 2004; Vergis et al., 2014). As discussed later (Section 1.3), there is some controversy regarding the anatomy associated with AOS. In general, brain damage that predicts AOS includes left hemisphere inferior frontal regions (Hillis et al., 2004; Richardson, Fillmore, Rorden, Lapointe, & Fridriksson, 2012); the insula (Dronkers, 1996; Dronkers & Ogar, 2004; Dronkers, Ogar, Willock, & Wilkins, 2004); motor, premotor, and supplementary motor areas (Basilakos et al., 2015; Graff-Radford et al., 2014; Josephs et al., 2012; Whitwell, Duffy, Strand, Xia, et al., 2013), as well as sensorimotor cortical areas in the post-central gyrus (Basilakos et al., 2015; Hickok, 2014a; Hillis et al., 2004).

1.2.1 The Dysarthrias. Dysarthria is a disorder of speech execution and control that can affect speech production at the levels of respiration, phonation, articulation, resonance, and prosody (Duffy, 1998). Dysarthria can be caused by a number of

neurologic impairments, but for the purpose of this study, only types most commonly associated with stroke will be discussed. These include upper and lower motor neuron dysarthria (spastic and flaccid dysarthria, respectively) and ataxic dysarthria (caused by damage to the cerebellum).¹ Dysarthria associated with motor neurons can result from an inability of the upper motor neurons to carry motor commands to the lower motor neurons (spastic, or upper motor neuron dysarthria), or an impairment in the lower motor neurons' ability to send these commands to the muscles they innervate (flaccid, or lower motor neuron dysarthria). Speech production in spastic dysarthria is perceptually "harsh" sounding, with a "strained and strangled" vocal quality (Duffy, 2005). Speech may be hypernasal and characterized by imprecise articulation (McNeil et al., 1997). In contrast, muscle atrophy, weakness, and/or paralysis cause speech impairments in flaccid dysarthria. Because speech musculature is innervated by lower motor neurons, problems with the neuron itself, the neuromuscular junction, or vocal tract musculature can lead to speech that is slurred, with reduced articulatory precision. Lastly, ataxic dysarthria can occur following cerebellar stroke. Its most striking impairment is reduced coordination between respiration and phonation, as well as decreased articulatory precision, altered prosody, and articulatory "breakdowns" (Kent, Kent, Duffy, et al., 2000).

Aspects of all three of the aforementioned dysarthria profiles sound perceptually similar when compared to AOS, especially with regard to imprecise articulation. However, the dysarthrias are more often accompanied by paralysis or paresis (e.g., flaccid and spastic types), and in the case of ataxic dysarthria, gross motor movements

¹ Hypo- and hyperkinetic dysarthrias can occur following damage to the basal ganglia, but these types are more common in degenerative processes (e.g., Parkinson's disease and Huntington's disease).

are also similarly uncoordinated, facilitating the diagnosis of ataxia. Table 1.1 presents similarities and differences between AOS and dysarthria. In the following sections, studies that have attempted to differentiate MSDs and aphasia will be reviewed. Less emphasis will be placed on dysarthria, as chronic dysarthria following unilateral cortical damage is rather uncommon (Duffy, 2005)².

Table 1.1.

Speech production impairments in AOS and dysarthria

AOS	Dysarthria	Both AOS and Dysarthria
-Distortion errors that increase with utterance length and/or complexity -Distorted sound substitutions/additions	-Muscle paralysis or paresis -Altered vocal quality -Decreased respiratory capacity is possible -Impairment evident across speech and non-speech tasks	-Distortion errors -Increased intersegment duration between sounds, syllables, and words (considered prosodic abnormalities) -Slow rate of speech -Reduced phrase lengths

1.2.2 Distinguishing AOS from Aphasia. AOS rarely occurs in isolation when it is caused by stroke. Accordingly, differential diagnosis between AOS and aphasia is difficult because of the co-occurrence of these disorders and the fact that both share similar behavioral profiles (i.e., false starts/restarts, groping and struggle for speech production, effortful articulation, shortened phrase lengths; Basilakos et al., 2015; Haley et al., 2013; Haley et al., 2012; Haley & Martin, 2011; McNeil et al., 1997; Strand et al., 2014; van der Merwe, 1997). Broca's aphasia is the most common concomitant language impairment (Duffy, 2005; Graff-Radford et al., 2014), as both AOS and Broca's aphasia

² Unilateral upper motor neuron (UUMN) dysarthria is argued to be a valid clinical diagnosis, but in prevalence estimates from visits to the Mayo Clinic speech pathology department, UUMN dysarthria only constitutes a small portion of MSD diagnoses (<8%; Duffy, 2005).

result from similar patterns of brain damage (Hillis et al., 2004; Richardson et al., 2012). In addition, certain aspects of conduction aphasia are difficult to discriminate from AOS, although the two disorders are less likely to co-occur. The following sections will review the challenges of distinguishing AOS from Broca's and conduction aphasia – two aphasia types with characteristic behaviors that have posed difficulties to the differential diagnosis of AOS.

1.2.3 Distinguishing AOS and Broca's Aphasia. Historically, AOS has been referred to by a number of terms that associate it with Broca's aphasia (e.g., *afferent motor aphasia*, *little Broca's aphasia*, and *expressive aphasia*, among other terms; see Duffy, 2005 and Ogar et al., 2005 for review and discussion). It has been suggested that AOS is simply a symptom of Broca's aphasia, and in fact, some diagnostic criteria for Broca's aphasia include AOS (see Ogar et al., 2005). However, cases in which Broca's aphasia occurs without AOS, or rare cases where AOS occurs in isolation, provide evidence suggesting that these disorders can be dissociated (Duffy, 2006; Duffy, Peach, & Strand, 2007; Duffy et al., 2015; Graff-Radford et al., 2014; Josephs et al., 2012; Moser et al., 2016; Richardson et al., 2012; Strand et al., 2014; Whitwell, Duffy, Strand, Xia, et al., 2013).

1.2.4 Distinguishing AOS and Conduction Aphasia. AOS and conduction aphasia can be difficult to differentiate perceptually due to the sound level errors that characterize production in both disorders. Conduction aphasia manifests as difficulty with word repetition and naming, and impairments in phonological processing are a common hallmark of conduction aphasia (Nadeau, 2000). Although both conduction aphasia and AOS are characterized by sound level errors, further inspection of errors that occur in

each disorder suggests that they arise from different production levels, with conduction aphasia dominated by errors at the level of phonemic encoding, and AOS dominated by phonetic-motoric processes (Itoh & Sasanuma, 1984; McNeil et al., 2000; McNeil et al., 1997; Rogers, 1997; Seddoh et al., 1996; van der Merwe, 1997). This distinction can be illustrated by inspecting the types of errors made by those with AOS and conduction aphasia. For example, individuals with AOS are often aware of their speech production errors (suggesting a relatively intact conceptual/phonological representation), but production errors are distorted and articulation is labored. Those with AOS often fail at self-correction, even after multiple attempts at production (van der Merwe, 1997; Wertz, LaPointe, & Rosenbek, 1984), but the presence of distortion errors suggests that unsuccessful self-corrections can be explained by disruptions in motor plans/programs (van der Merwe, 1997).

In contrast, individuals with conduction aphasia tend to be aware of their errors, but attempts at self-correction are characterized by variable phonemic approximations of the target, and importantly, phonemes tend to be well articulated. This characteristic has been referred to as *conduite d'approche*, or “phonemic approximation” (Goodglass, 1992). It has been suggested that the frequency of (well-articulated) phonemic errors can be explained by higher-level difficulties in the sensorimotor transform of auditory targets (corresponding to target phonemes) into a form that can be handled by the motor system (Buchsbaum et al., 2011).

The distinction between *conduite d'approche* and articulation errors in AOS lies in the inability of the “auditory representations” of speech to guide lower-level production processes (conduction aphasia), rather than an impairment in articulatory

centers themselves (AOS; Buchsbaum et al., 2011; Hickok et al., 2014). To reiterate, phonemic paraphasias that occur in *conduite d'approche* are often mistaken for articulatory errors in AOS (Haley et al., 2013; Haley et al., 2012), but are due to higher-level deficits in phonological processing, not deficits with articulatory planning (Itoh & Sasanuma, 1984; McNeil et al., 1997; Rosenbek, Kent, et al., 1984; van der Merwe, 1997). Table 1.2 presents sound level impairments that occur in aphasia, AOS, and both aphasia and AOS.

Table 1.2.

Speech production impairments in AOS and aphasia

AOS	Aphasia	Both AOS and Aphasia
-Distorted substitution and addition errors -Prosodic disturbances	- <i>Conduite d'approche</i> (phonemic approximations; conduction aphasia) -Sound level errors due to breakdowns in phonological encoding -Articulation is not distorted	-False starts/restarts -Decreased fluency and rate of speech

1.3 NEUROANATOMICAL CORRELATES OF AOS AND APHASIA

A large body of research has attempted to differentiate AOS from the aphasias behaviorally (Cunningham et al., 2015; Galluzzi et al., 2015; Haley et al., 2013; Haley et al., 2012; Haley & Martin, 2011; Itoh & Sasanuma, 1984; Jacks, 2008; Maas, Mailend, & Guenther, 2015; Maas, Robin, Wright, & Ballard, 2008; Seddoh et al., 1996; Ziegler & von Cramon, 1985, 1986a, 1986b), neuroanatomically (Ballard, Tourville, & Robin, 2014; Basilakos et al., 2015; Dronkers, 1996; Graff-Radford et al., 2014; Josephs et al., 2006; Ogar et al., 2006; Whitwell, Duffy, Strand, Xia, et al., 2013), and theoretically

(Ballard, Granier, & Robin, 2000; Ballard et al., 2014; Duffy & K. A. Josephs, 2012; Strand et al., 2014; Ziegler, 2002; Ziegler et al., 2012). Yet, determining whether impaired production is caused by breakdowns in phonetic-motoric or phonemic processes remains difficult in perceptual speech evaluations (McNeil et al., 2000; Wambaugh, Wright, Nessler, & Mauszycki, 2014). To complicate matters, even though individuals with AOS demonstrate several hallmark behaviors (reduced speech rate, atypical prosody, speech sound distortions, articulatory groping), not all those with AOS demonstrate all of these behaviors to the same extent (e.g., Cunningham et al., 2015; Galluzzi et al., 2015; Maas et al., 2015). In addition, these syndromes are not mutually exclusive and while there are clear dissociations, many individuals suffer from several of these impairments. Furthermore, the one currently available published test for AOS (*ABA-2*; Dabul, 2000) suffers from poor validity, as it can falsely classify individuals with aphasia as having AOS (McNeil et al., 2004).

Neuroimaging results from patients with AOS are also inconsistent across studies. Most imaging studies have been conducted with individuals at the chronic phase of stroke (greater than six months post-onset), and have incorporated either lesion overlap methods (e.g., Dronkers, 1996; Ogar et al., 2006), voxel-based correlation methods (Baldo, Wilkins, Ogar, Willock, & Dronkers, 2011; Basilakos et al., 2015; Hickok et al., 2014; Hillis et al., 2004; Richardson et al., 2012; Trupe et al., 2013), or qualitative analysis of individual cases (Graff-Radford et al., 2014; Moser et al., 2016). One of the most well-known, seminal studies to relate AOS to brain damage was by Dronkers (1996), who found that a group of patients with chronic AOS demonstrated 100% lesion overlap in the superior precentral gyrus of the insula (SPGI). Dronkers' was not the first to suggest that

the insula was important for speech production. Historical accounts from a 1908 meeting of the Society of Neurology of Paris report debates surrounding the insula (Cole & Cole, 1971), but discussion was generally unfavorable or inconclusive regarding the insula's role in speech production (Cole & Cole, 1971; Ojemann & Whitaker, 1978).

Nevertheless, for years after Dronkers' (1996) study, the SPGI, or more generally, the anterior insula, was considered to be the region primarily responsible for AOS, as findings from later group studies (Baldo et al., 2011; Dronkers, 1996; Dronkers et al., 2004; Ogar et al., 2005; Ogar et al., 2006; Wise, Greene, Büchel, & Scott, 1999) and one individual case of isolated insula infarct (Nagao, Takeda, Komori, Isozaki, & Hirai, 1999) concluded that the insula is implicated in post-stroke AOS.

Later studies that incorporated voxel-based lesion symptom mapping (VLSM), did not replicate Dronkers and colleagues' findings. Instead, these studies showed that the inferior frontal gyrus *pars opercularis* (IFGpo) was most predictive of AOS, not the insula (Hillis et al., 2004; Richardson et al., 2012). In further contradiction to Dronkers' seminal (1996) findings and those suggesting the IFGpo is the neural correlate of AOS, more recent studies have found that premotor, supplementary motor, and sensorimotor areas are predictive of AOS (Basilakos et al., 2015; Graff-Radford et al., 2014; Hickok et al., 2014; Josephs et al., 2012; Josephs et al., 2006; Whitwell, Duffy, Strand, Xia, et al., 2013). None of these more recent studies found that damage to the insula is correlated with AOS, but some have replicated findings implicating the IFGpo (Graff-Radford et al., 2014; Hickok et al., 2014). All together, these studies have made a case for different regions being crucially implicated in AOS – the insula, the IFGpo, and sensorimotor

cortical areas. However, we suggest that these discrepant findings across studies might be explained by differences in neuroimaging analysis methods or the different, often subjectively determined, AOS diagnoses used in each study. These issues will be discussed in the following sections.

The above-mentioned studies utilized different imaging analysis techniques – either lesion-overlap methods (Dronkers, 1996; Ogar et al., 2006) or VLSM (Basilakos et al., 2015; Duffy & Josephs, 2012; Graff-Radford et al., 2014; Hickok et al., 2014; Hillis et al., 2004; Josephs et al., 2012; Whitwell, Duffy, Strand, Xia, et al., 2013). Lesion overlap methods, while informative, can be heavily influenced by *a priori* subject selection criteria and brain anatomy itself (Hillis et al., 2004; Rorden & Karnath, 2004). With regard to Dronkers' (1996) findings, she reported that the insula was the only brain region to predict perfectly AOS: all patients with insula damage had AOS; there was never a case where a patient had insula damage but did not have AOS. Although Dronkers' (1996) findings argued strongly in support of the insula's role in AOS, it does not rule out the possibility that a patient can have AOS without insula damage, or that a patient can have insula damage and not AOS.

With regard to brain anatomy, the insula is commonly affected following large MCA strokes, especially when the IFG is also compromised (Finley et al., 2003; Hillis et al., 2004). Furthermore, structural methods utilized at the chronic stage (e.g., lesion overlap and VLSM) do not take into account functional changes due to diaschisis or reorganization. To avoid some of these issues, Hillis and colleagues (2004) tested 80 patients at the acute phase of stroke (within 24 hours of onset) using diffusion weighted and perfusion weighted imaging (DWI and PWI, respectively) to determine neural

correlates of AOS in acute stroke. Hillis et al. (2004) found that of the patients with AOS, 14 had compromise (either decreased perfusion and/or infarct) to the left anterior insula, but 26 demonstrated either structural damage or hypoperfusion to Broca's area. Of these cases, there was only one case where the insula was compromised and the IFGpo was intact, but there were 13 cases where the insula was normal and the IFGpo was affected. Additionally, there were five patients with AOS who did not demonstrate damage to either the anterior insula or Broca's area; instead, these cases had hypoperfusion or structural damage to the pre- or post-central gyri. Hillis et al. (2004) concluded that not all cases of acute AOS have hypoperfusion or infarct to the insula, and not all cases of hypoperfusion or infarct to the insula lead to AOS. According to this view, Dronkers' findings with chronic patients may reflect that the structural scans (which were relatively crude by modern standards) may have consistently failed to detect disruption of the IFGpo and sensorimotor cortical areas.

The discordant findings between Dronkers (1996) and Hillis et al. (2004) could be attributed to the different analysis techniques employed, or that those with chronic AOS have different patterns of damage, leading to deficits that are less amenable to reorganization or treatment. To address some of these concerns, Richardson et al. (2012) analyzed lesion damage using both lesion overlap and VLSM methods in a sample of chronic-post stroke individuals, using the same diagnostic criteria as Dronkers (1996). One group of patients presented with AOS and aphasia, and the other presented with aphasia only. The results of each analysis yielded very different findings. Results from the lesion overlap method showed that the insula was the greatest area of overlap, while results from the VLSM revealed that the IFGpo was most predictive of AOS.

Taken together, Richardson et al.'s results suggest that discrepancies between the aforementioned studies can be explained by the use of different analysis techniques. For example, work by Dronkers (Dronkers, 1996; Dronkers & Ogar, 2004; Dronkers et al., 2004; Ogar et al., 2006) and Richardson et al. (2012) classified patients using criteria established by Wertz (1984), but when the same diagnostic criteria were used in the same group of patients, results differed greatly based on the analysis method utilized (i.e., VLSM *vs.* lesion overlap). As discussed above, lesion overlap methods are subject to biases in subject selection and neuroanatomy (Hillis et al., 2004; Rorden & Karnath, 2004), meaning the differences between these two studies can most likely be attributed to the analysis methods used.

It should also be acknowledged that the lack of consistent findings across studies can be explained by the fact that many diagnostic measures are subjective, meaning investigators may not rate all measures similarly. More recent studies (Basilakos et al., 2015; Graff-Radford et al., 2014; Josephs & Duffy, 2008; Josephs et al., 2012; Josephs et al., 2006; Whitwell, Duffy, Strand, Xia, et al., 2013) have used the Apraxia of Speech Rating Scale (ASRS; Strand et al., 2014) criteria for classifying AOS. The ASRS is a 16-item scale that was developed by clinicians at the Mayo Clinic (Rochester, MN) who have extensive experience with motor speech disorders, and who developed this scale based on expert consensus. The difference between the Wertz (1984) criteria and the ASRS is that the ASRS delineates behaviors that occur exclusively in AOS from those that can occur in AOS and aphasia, AOS and dysarthria, and between all three disorders. Interestingly, recent studies that have implemented the ASRS in lesion-based or positron emission tomography (PET) scan studies found that AOS is associated with damage to

cortical motor areas (Basilakos et al., 2015; Duffy, 2006; Itoh & Sasanuma, 1984; Jordan & Hillis, 2006; Josephs et al., 2012; Josephs et al., 2006; Whitwell, Duffy, Strand, Xia, et al., 2013), providing preliminary evidence that use of the same diagnostic measure, shown to have good validity and reliability (Strand et al., 2014) may help minimize inconsistent results across studies.

Finally, discrepant findings from previous studies could be explained by anecdotal evidence that different “types” of AOS exist. That is, it has been suggested that patients with AOS can be separated into groups based on whether speech production errors are characterized predominately by articulatory distortion errors, or atypical prosody (Duffy, Strand, Whitwell, & Josephs, 2013; Josephs et al., 2013). In a study of 28 patients with primary progressive AOS (PPAOS) and progressive AOS and aphasia, Duffy and colleagues (2013) reported that six of the patients produced speech that was heavily marked by distortion errors, while 18 of the patients had fewer distortion errors but demonstrated atypical prosody (i.e., syllable segmentation, lengthened intersegment durations). The remaining four patients could not be assigned a type. Inspection of these patients’ neuroimaging data suggested that these types might be caused by slightly different patterns of neurodegeneration, concluding that different regions may support the planning/programming of different units, and that these processes may not be affected uniformly when speech planning/programming processes are compromised.

The existence of AOS types has been speculated in post-stroke AOS (Ballard, Barlow, & Robin, 2001; Croot, 2002; Duffy & Josephs, 2012; Rosenbek, Kent, et al., 1984). If different types of post-stroke AOS exist, providing empirical evidence regarding whether or not distortion errors and prosodic impairments are represented by

different patterns of brain damage would be beneficial to the study of AOS. Future studies may need to implement measures that reliably capture characteristic of each type, as collapsing data from patients with different AOS types into a single group may obscure subtle behavioral differences that could have important implications for understanding levels of speech production affected, as well as the brain regions that are responsible for these processes. Taken together, this discussion highlights the need for objective measures of speech planning/programming that can quantify rhythmic/prosodic production, in addition to those that measure articulatory precision at the phonemic and/or featural level. Regardless of whether or not post-stroke AOS “subtypes” indeed exist, such measures may more accurately quantify apraxic impairments, and the extent that production processes are disrupted following a left hemisphere stroke.

These aforementioned issues are of importance to the study of post-stroke aphasia and AOS from a theoretical and clinical standpoint. The current state of AOS research emphasizes that using diagnostic labels when diagnostic criteria are not agreed upon can yield results that are not generalizable across studies, limiting the extent to which studies can be replicated and contribute meaningfully to the AOS literature. Theoretically, the lack of consistent findings across studies has obscured full understanding of the nature of AOS, especially with regard to how the stages of speech production processes are affected differently in AOS and aphasia. Differences in results across studies have also limited the localization of the neural correlates of AOS and motor speech planning/programming in general.

Clinically, the difficulty in differentiating the underlying cause of production impairments poses a challenge during the treatment planning process. In-depth evaluation

of speech and language is often necessary to determine if production impairment is dominated by AOS or aphasia (McNeil et al., 2000; Wambaugh et al., 2014), as quick and reliable measures to differentiate aphasia from motor speech disorders are only beginning to emerge (e.g., Strand et al., 2014) but are in need of further validation. Unfortunately, many clinicians have time only for a cursory assessment in the acute setting, and patients seeking rehabilitation in the chronic phase of stroke are often limited to one evaluation session due to restrictions by third-party payers. Accordingly, some argue that treatment of aphasia should be prioritized, and that treating expressive language in general may also benefit speech production (Dabul, 2000). However, others argue that AOS is not amenable to language-based treatments because many treatments for language production do not provide patients sufficient opportunities to target articulation, nor do they address production errors from a motor control standpoint (Ballard et al., 2015; McNeil et al., 2000; Wambaugh et al., 2006; Wambaugh, Kalinyak-Fliszar, West, & Doyle, 1998; Wambaugh, Nessler, Cameron, & Mauszycki, 2012; Wambaugh, Nessler, Cameron, & Mauszycki, 2013; Wambaugh et al., 2014).

As evidenced by the above discussion, the study of post-stroke speech and language disorders would benefit from validation of objective measures. Refining such measures for use with post-stroke patients can aid clinicians during diagnostic processes, facilitating the differentiation of AOS from aphasia. This could dramatically improve clinical care, both by offering improved prognostic expectations and identifying bespoke treatment options. Accurate measures could also help validate new treatments by better counterbalancing patients at inclusion and more accurately measuring the influence of the interventions. Improving upon diagnostic measures also has far reaching theoretical

implications. For the purpose of further research, adopting behavioral measures that can be replicated reliably across studies will limit the potential that differences in subjective diagnosis of speech production impairments affect study outcomes. Ultimately, more consistent findings across studies can help rectify debates regarding AOS, including its theoretical nature and behavioral manifestations. As for now, theories that account for AOS must include how motor speech production processes can break down, and should do so within the context of other related production processes, i.e., word finding, phonemic encoding, syntactic computations, as well as motor speech execution, internal monitoring and feedback control. Unfortunately, a great deal of evidence is lacking to fully explain AOS in this context.

1.4 MODELS OF SPEECH PRODUCTION

This section is devoted to the discussion of theoretical bases of speech production, focusing first on normal production, followed by disordered production. This section highlights how contemporary production models, although still in development, have potential explanatory power for AOS in the broader context of speech and language processes.

Speech production involves a complex, well-orchestrated sequence of movements, executed to match auditory goals corresponding to the sounds that compose words in a given language (Guenther, 1994, 1995; Guenther, Hampson, & Johnson, 1998; Guenther & Vladusich, 2012; Hickok, 2012a, 2014a; Hickok & Poeppel, 2004; Levelt, Roelofs, & Meyer, 1999; Perkell et al., 2004; Tourville & Guenther, 2011). A number of models from various disciplines (i.e., psychology, linguistics, neuroscience, motor

control) have been proposed to explain the processes that occur from conceptualization to lexical selection, phonemic encoding, and finally articulation (Hickok, 2014a; Hickok & Poeppel, 2000; Levelt, 1983, 1992; Levelt et al., 1999; Ziegler et al., 2012). Perhaps one of the most influential models today is that of Levelt and colleagues (Levelt et al., 1999).

Levelt and colleagues' model explains production in essentially three serial stages. First, production begins with the conceptual level, motivated by the goals and intentions of the speaker (1999). At this stage, speakers have access to a network of lexical concepts. Next, lexical selection occurs, where a word (referred to here as a "lemma") is chosen from the lexicon. Lexical selection completes the semantic stages of production and is followed by phonemic encoding and processes to prepare for impending articulation. Levelt and colleagues propose the existence of a mental syllabary, where gestures (movements) that correspond to frequently produced syllables, words, and even short phrases are stored for easy access (Levelt et al., 1999; Levelt & Wheeldon, 1994). As articulatory gestures are accessed from the syllabary, syllabification occurs extemporaneously, during the actual process of production, as stress and intonation are partially dependent on the context of the word within its utterance (Levelt, 1999; Levelt et al., 1999).

Levelt et al.'s model (1999) does not provide a fine-grained explanation of neuroanatomical structures that support speech production processes. Nevertheless, the work by Levelt and colleagues (Levelt, 1983, 1992, 1999; Levelt et al., 1999) has laid the groundwork for more contemporary models; for example, the Directions into the Velocities of the Articulators (DIVA; Guenther & Vladusich, 2012; Tourville & Guenther, 2011) and more recently, the State Feedback Control (Houde & Nagarajan,

2011) and the Hierarchical State Feedback Control (HSFC; Hickok, 2012a, 2014a) models.

The DIVA model (Guenther, 1994, 1995; Guenther & Vladusich, 2012; Tourville & Guenther, 2011) picks up where Levelt and colleagues' model leaves off; at the level of the syllable and lower articulatory stages (Guenther, 2006). The focus of the DIVA model is on sensorimotor processes that support speech production, emphasizing the role of feedforward and feedback control of articulation (Golfinopoulos et al., 2011; Golfinopoulos, Tourville, & Guenther, 2010; Guenther, 1994; Guenther, 1995; Guenther, 1995; Guenther, 2006; Guenther, Ghosh, & Tourville, 2006; Tourville & Guenther, 2011; Tourville, Reilly, & Guenther, 2008). The DIVA model's account of speech production begins from a developmental perspective. That is, in early pre-lingual development, input gained from auditory feedback of an infant's babbling is used to develop and refine oral motor movements that match auditory targets of the spoken language. Over time, correspondence between language-specific speech sounds, articulatory movements and the somatosensory (auditory and sensorimotor) feedback are learned. This information is stored in *speech sound map cells*, similar to Levelt et al.'s (1999) "syllabary," for quick access during production (Guenther, 1995, 2006; Guenther et al., 2006).

Although feedback control is important in maintaining fluent production, solely relying on feedback monitoring is not sufficient due to the fact that speech production occurs rapidly. Sensory feedback must be processed before this information can be utilized by the production system, but time lags in the neural transmission of feedback are incurred due to processing constraints at the level of axons and synapses (Kandel, Schwartz, & Jesell, 2000). Furthermore, the auditory cortex must process incoming

feedback, and it has been estimated that in some cases, the processing of acoustic feedback can require between 30-100 ms before this information can be utilized (Houde & Nagarajan, 2011). To account for these delays in feedback transmission, auditory and sensorimotor input guides speech production to tune internal, forward models as development continues. These internal, feedforward models learn motor commands associated with speech sounds and can initiate production processes without relying on acoustic and sensorimotor input. Importantly, feedforward commands anticipate the articulatory trajectories and resulting auditory feedback that corresponds to speech motor movements. In the case that internally predicted outcomes do not match the externally predicted production targets, corrective commands can be sent to remedy these errors before they are realized during ongoing production (Hickok, 2012b; Houde & Nagarajan, 2011; Tourville et al., 2008). Internal feedforward commands are integral in speech production, as once learned, production can proceed with little input from auditory and sensorimotor feedback unless some form of perturbation is present, i.e., a bite block, loss of hearing acuity (Behroozmand, Karvelis, Liu, & Larson, 2009; Greenlee et al., 2013; Guenther, 1995, 2006; Guenther et al., 2006; Hickok, 2012b; Houde & Nagarajan, 2011; Jacks, 2008; Tourville et al., 2008).

The DIVA model is one of the first extensively tested models of speech production to explain feedforward and feedback control over the course of development, normal production, and in the case of impairment to the speech or hearing systems. More recently developed models have built upon DIVA by incorporating principles of limb motor control to motor speech production. These models include the SFC model (Hickok, Houde, & Rong, 2011) and a more recent iteration, the HSFC model (Hickok, 2012a).

SFC models are similar to DIVA in that they propose the end product of speech production is to reach an auditory target that will be understood by listeners. However, SFC models differ from DIVA in that they account for the exact position and velocity of the articulators during speech production so that when an error (either auditory or somatosensory) is detected, corrective commands take into account articulator movement trajectories so these commands do not impede with ongoing production (Houde & Nagarajan, 2011). Although SFC models offer a way to bridge limb motor control theories to the study of speech production, neither the SFC nor the DIVA models explain the role of the linguistic-conceptual system in the monitoring of ongoing production processes.

Accordingly, the HSFC model was developed to account for the role of the conceptual-linguistic system in production and monitoring processes. The HSFC model further expands on SFC models to explain speech production as occurring in two "hierarchical" steps: a higher level that identifies a sensory target (stored auditory representation) and a lower level that is responsible for programming articulatory motor movements for this target. When a lexical item is selected, its corresponding sensory target guides lower-level phonetic-motoric processes. Both processes are enacted in parallel, but are mediated by different neuroanatomical regions and guided by feedback from different sources, as detailed below.

Following lexical selection, auditory representations of the lexical item that are stored in the superior temporal gyrus (STG) and superior temporal sulcus (STS) are sent to area Spt (the posterior Sylvian fissure at the boundary of the parietal and temporal lobes) to “transform” auditory targets to motor targets. These motor targets are then sent

to the IFGpo (Brodmann Area, BA, 44) for syllable-level planning. Auditory feedback guides this level of production. At the lower phonetic-motoric level, production targets are selected in sensorimotor regions and then sent to cortical motor areas (BA6 and M1), mediated by cerebellar guidance for somatosensory processing and sensorimotor feedback. Accordingly, the HSFC proposes that speech production is guided by the conceptual representation of a word, its corresponding auditory-syllabic and phonetic-motoric forms, and afferent input from auditory and sensorimotor feedback. The HSFC is still in refinement, however.

Relevant to the study of AOS, some theoretical accounts suggest that speech planning can break down due to impaired access to stored syllabic units (Ballard et al., 2001; Varley & Whiteside, 2001; Varley, Whiteside, & Luff, 1999). Others do not support this contention; rather, they provide evidence that AOS results from the failure to retrieve motor programs for units as small as gestural and phonemic, to those as large as syllabic and metrical, depending on various aspects of the target word (Aichert & Ziegler, 2004; Mailend & Maas, 2013; Ziegler, 2005, 2009). When considered in the context of the HSFC, breakdowns in production could theoretically occur at the syllabic and/or phonemic levels. It has been speculated that production and feedback monitoring at these levels can be affected differently in individuals with AOS, meaning production errors that occur in AOS can be dominated by syllable level (prosodic) or phonetic level (distortion) errors (Maas et al., 2015). As such, the HSFC may offer some explanations for these differences.

To date, only a handful of studies have used the HSFC as a theoretical model to interpret findings pertaining to AOS (e.g., Basilakos et al., 2015; Maas et al., 2015). To

our knowledge there are no published studies that have explicitly investigated whether the HSFC can indeed provide a theoretical account of apraxic errors, or a framework to distinguish between impaired processes that lend themselves to errors that occur in AOS, from those that can occur in aphasia. However, there is preliminary support for the HSFC's ability to explain disordered production from studies that have investigated the neural correlates of phonemic errors in conduction aphasia (Buchsbaum et al., 2011; Okada & Hickok, 2006; Ueno & Lambon Ralph, 2013). Nevertheless, because the HSFC model describes production at both the auditory-syllable and phonetic-motoric levels, in addition to providing an explanation for the role of feedforward and feedback processes during speech production, this model may be instrumental in explaining phonemic and phonetic level errors in aphasia and AOS.

1.5 CHAPTER SUMMARY

Publications pertaining to AOS that do not include discussion of its controversies are almost nonexistent, as the study of AOS has been shadowed by debates regarding its theoretical nature and even the extent that it is its own unique clinical entity (Itoh & Sasanuma, 1984; Martin, 1974; McNeil & Kent, 1990; McNeil et al., 1997; Rosenbek, Kent, et al., 1984). As detailed in this chapter, various factors could explain discrepant findings regarding the localization of AOS. These factors include a) differences in methods used to analyze lesion damage, and b) the lack of uniformity in behavioral definitions adopted to classify patients with different disorders. In addition, research into brain-behavior relationships has primarily focused on regional damage, and little is known about how compromise to regional and network result in apraxic impairment.

The research reviewed here is by no means exhaustive, but it has contributed to a body of literature motivated to improve the diagnostic criteria for AOS. This work has greatly improved our understanding of AOS, but the study of AOS remains a challenge. That is, investigations into AOS would be made easier by systematic study of cases of pure AOS without concomitant aphasia. However, AOS rarely occurs without concomitant aphasia, making it difficult to study production errors that are specific to AOS. Furthermore, classifying behaviors into the categories of 1) phonetic/motoric in nature (i.e., corresponding to AOS) or 2) phonemic (i.e., corresponding to aphasia) is difficult due to perceptual limitations in making such classifications (Itoh & Sasanuma, 1984), and the availability of very few objective measures (Ballard et al., 2016; Vergis et al., 2014). Finally, studies in which diagnostic classifications are defined *a priori* for between groups analyses may suffer from circularity confounds. By studying the behaviors of a group of individuals who have been assigned a diagnostic label, it is possible that the behaviors under examination are subject to bias that is inherent in identifying the patient group in the first place, especially in the case of diagnosing AOS (Cunningham et al., 2015; Galluzzi et al., 2015).

1.6 STUDY PURPOSE

The purpose of this study was to address some of the challenges in the diagnosis of AOS that were discussed in the previous sections. To this end, this study aimed to determine the extent that objective acoustic measures can complement currently available measures of post-stroke speech production to improve differential diagnosis of AOS and aphasia. The aims are as follows: a) to improve differential diagnosis of post-stroke communication disorders by identifying objective measures that classify groups of

patients with similar production impairments (i.e., resulting from AOS), from those whose production impairments are different (i.e., phonemic level impairment, attributed to aphasia), and b) to use measures that are most predictive of group classification to determine how patterns of brain damage (both regional and network damage) gives rise to these impairments. Objective measures that have shown to be predictive of AOS classification (VOT, a measure of inter-articulatory planning and coordination; pairwise variability index, PVI, a measure of rhythm/prosody; and narrow phonetic transcription) were investigated along with subjective, perceptual measures (i.e., those included on the ASRS). Additionally, variables obtained from amplitude envelope modulation spectra, shown to facilitate differential diagnosis of the dysarthrias (Liss, LeGendre, & Lotto, 2010), were investigated for the first time in a large sample of individuals with AOS.

Because no “gold standard” for AOS classification exists, participants were grouped based on the extent that speech is characterized predominantly by a) suspected AOS, b) production errors that can be attributed to aphasia, or c) the absence of production errors beyond what could be considered within normal limits for typical speakers. Scores from the ASRS and the WAB were used to define these groups. According to Strand et al. (2014), participants with an ASRS total rating greater than or equal to eight (ASRS total ≥ 8), and at least one item rated from ASRS Items 1.1-1.6 (behaviors considered unique to AOS; scale presented in Appendix A) are likely to present with production errors consistent with AOS. Using Strand et al.'s (2014) recommendations, performance on the WAB was used to guide ASRS aphasia severity scoring. Participants with WAB Aphasia Quotient (WAB AQ) scores greater than 93.8 (the cutoff score for aphasia diagnosis), who also have ASRS total scores less than eight,

served as a "stroke control" group, i.e., a group of participants in the chronic stage of stroke, but with no chronic speech or language impairment measured by WAB or ASRS criteria.

The purpose of these groups was to obtain objective and subjective behavioral measures of phonemic, phonetic, and prosodic impairments in individuals with chronic stroke, and importantly, to characterize performance of those with AOS compared to individuals with aphasia or no speech/language impairment, but who have experienced a stroke. Supervised machine learning analyses were used to determine how these measures distinguish between participant groups. VLSM and connectome-symptom mapping were used to determine how different production errors relate to patterns of post-stroke brain injury.

1.7 SPECIFIC AIMS AND HYPOTHESES

1.7.1 Specific Aim 1. Determine if objective measures differentiate speakers with suspected AOS from speakers with other production impairments. Specifically, the purpose of this aim was to determine the extent that measures of phonemic, phonetic and prosodic impairment could account for between group differences in diagnostic classification.

Hypothesis 1a. Objective measures of phonetic impairment (e.g., voice onset time, distorted sound errors) will distinguish participant groups into those with suspected AOS (and concomitant aphasia) from those without AOS, but who may have production errors related to aphasia (i.e., phonemic errors).

Hypothesis 1b. Prosodic measures will also distinguish between speakers with phonetic and phonemic impairment (McNeil et al., 2004; Vergis et al., 2014), improving classification of those with AOS from those without AOS.

Hypothesis 1c. If post-stroke speakers with phonetic impairment can indeed be classified into “types,” a subgroup of participants with phonetic or prosodic impairments should emerge through behavioral analyses.

1.7.2 Specific Aim 2: Identify patterns of regional and network damage that correlate with the *behaviors* instrumental in predicting group assignment.

Hypothesis 2a: Measures sensitive to phonetic level impairments will be predicted by damage to motor and sensorimotor areas. Further inspection of network damage will reveal decreased connectivity between motor areas and regions involved in sound level (phonemic) production (e.g., decreased connection between inferior frontal areas and BA6/M1). Additionally, connections between cerebellar and motor areas will be implicated, due to the importance of these regions for sensorimotor control and internal monitoring for speech production processes (Hickok, 2012).

Hypothesis 2b: Measures sensitive to phonemic level impairments will be predicted by damage to posterior Spt areas (Buchsbaum et al., 2011; Okada & Hickok, 2006; Pa & Hickok, 2008), or anterior dorsal stream areas (Schwartz, Faseyitan, Kim, & Coslett, 2012). Connectivity between posterior temporoparietal areas and inferior frontal regions (i.e., IFGpo) will be reduced (Buchsbaum et al., 2011).

1.8 PILOT DATA

Exploratory analyses were conducted on speech and language test scores for 44 chronic post-stroke individuals using a) multidimensional scaling (MDS), b) a TwoStep

clustering approach (SPSS, version 22), and c) neuroimaging analysis. Participants included in this analysis were given a thorough battery of speech and language testing in which ASRS and WAB scores were obtained. Diagnosis of AOS was made based on ASRS guidelines (total ASRS score ≥ 8 at least one item rated from the list of behaviors that occur specifically due to AOS), and aphasia diagnoses were made based on WAB AQ scores. Participant diagnoses are as follows: two with AOS only, 13 with AOS and aphasia, 27 without AOS. Additional details regarding this sample can be found in Basilakos et al. (2015).

1.8.1 Classification results. Summary scores from the ASRS and WAB subtests (fluency, naming, spontaneous speech and repetition) were standardized and included in the MDS and TwoStep analyses. MDS calculates a distance measure for each of the input items to determine a configuration that reduces stress, a goodness of fit measure. Results from this analysis reveal that ASRS and WAB scores clustered into two dimensions. Stress was 0.095, which is considered “Good” (Kruskal & Wish, 1978). Calculated R-Square for this model was .98. Figure 1.1 (see next page) presents the Euclidean distances for each of the variables. It should be noted that the dimensions that result from an MDS analysis are arbitrary; therefore, further analysis of the dataset is necessary to determine which measure(s) are most important for group classification.

1.8.2 TwoStep Cluster Analysis. To determine which measures were most important for group classification, behaviors from the above MDS analysis were analyzed using the “TwoStep Clustering” procedures in SPSS (Version 22). It should be noted that scores from each of the 16 ASRS items were also included in this analysis (see Appendix A for a list of all 16 items), and the overall ASRS severity scores were removed, as the

purpose here was to determine specific *behaviors* that were most important for determining participant groups.

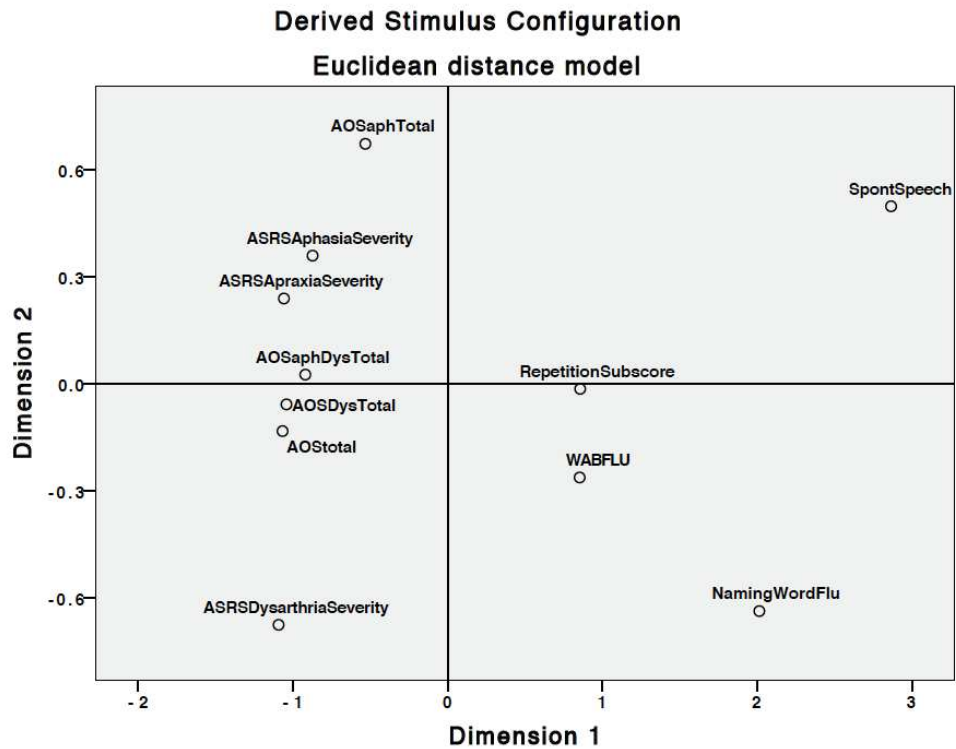


Figure 1.1. Results from preliminary MDS analysis. Items entered into this analysis include subtotals from the ASRS and select scores from the WAB (speech repetition, fluency, naming and word finding, and spontaneous speech).

The TwoStep cluster analysis begins with pre-clustering (step one) and grouping the data into sub-clusters (step two; see SPSS, 2001 for additional details). Step one begins clustering each successive input, using a measure of distance to form cluster feature (CF) trees. This part computes distance measures to iteratively compare whether the current input is similar to the prior input, and if so, it is placed in the same leaf within the same CF tree. If it is not, then a new leaf is formed. This process continues until all inputs have been assigned a leaf. In step two, clusters are formed using hierarchical clustering methods and a bootstrap method to determine the set of measures that

minimize variance. The number of desired groups can be specified *a priori*; however, for this analysis, the default number of maximum groupings was selected, i.e., between one-25, meaning up to 25 groupings could have emerged from this analysis. At the completion of step two, a group assignment is given to each participant entered in the analysis.

When scores from each ASRS item were included in the TwoStep analysis, along with WAB fluency, spontaneous speech, naming/word finding, and overall aphasia severity (WAB AQ) scores, two groups emerged. Item 1.3 on the ASRS (“Increased distortions/distorted sound substitutions with increased utterance length or increased syllable/word articulatory complexity”) emerged as the most important factor in predicting group membership. This factor is one of the six ASRS items that is unique to AOS. The next two items of importance included Item 2.6 (“Increased intersegment durations,” a feature shared between AOS and dysarthria), and Item 1.1 (“Distorted sound substitutions,” a feature unique to AOS). A rank order of predictor importance is presented in Figure 1.2.

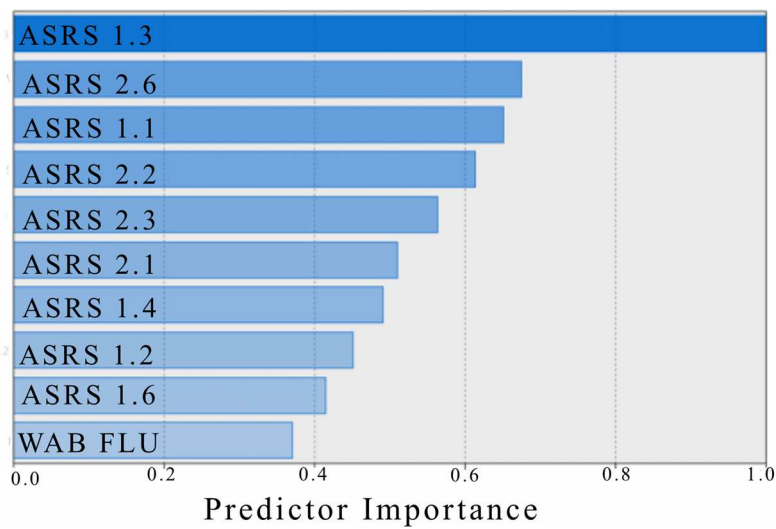


Figure 1.2. Rankings of predictor importance from the TwoStep cluster analysis. Note: Predictor Importance values (x-axis) are arbitrarily defined by the TwoStep procedure.

Next, cluster assignment obtained for each participant entered into the analysis was examined further to determine the extent that the predicted cluster assignment agreed with ASRS AOS diagnosis. For the 15 individuals who received a diagnosis of AOS, the TwoStep procedure assigned all but two individuals to the same group. For the 27 individuals without AOS, all of these members were classified together. Table 1.3 presents diagnostic accuracy for both groups.

Table 1.3

Group classification results from the TwoStep cluster analysis

Classification Based on TwoStep Clustering Results			
Clinician Classification	Group 1	Group 2	Totals
AOS	13	2	15
No AOS	0	27	27

1.8.3 Imaging Analysis. The top three items that emerged from the TwoStep analysis were analyzed in a series of univariate VLSM analyses (using nii_stat software <http://www.nitrc.org/projects/niistat/>) with 4000 permutations to control for multiple comparisons. The purpose of the neuroimaging analyses was to determine the extent that speech production errors could be dissociated neuroanatomically. The results revealed that the different behaviors were predicted by patterns of brain damage to the precentral gyrus, supramarginal gyrus (SMG), and superior longitudinal fasciculus (SLF), as well as additional regions presented in Table 1.4 and Figure 1.3. These results replicate findings of cortical motor areas implicated in AOS. However, the role of post-central areas, such

as the SMG, is not commonly reported in contemporary literature (for exceptions, see Hickok et al., 2014; Basilakos et al., 2015).

Table 1.4.

Results from preliminary univariate VLSM analyses

The percentage (in parentheses) reports proportional damage to each region implicated in the three behaviors presented. A list of abbreviations can be found on Pages viii-x.

Distorted Additions/ Substitutions	Increased Intersegmental Durations	Distorted Substitutions
SMG (45%)	SMG (41.2%)	PrCG (63.5%)
SLF (12.3%)	SLF (12.1%)	SMG (6.5%)
pSTG (9.7%)	pIns (8.8%)	PoCG (6%)
STG (8.7%)	pSTG (8.8%)	AG (5.8%)
PrCG (8 %)	Ex Capsule (7.7%)	STG (5.5%)
PoCG (6.5%)	STG (7.5%)	pMTG (5.5%)
pIns (5.8%)		SPG (5.4%)
		MFG (5.2%)

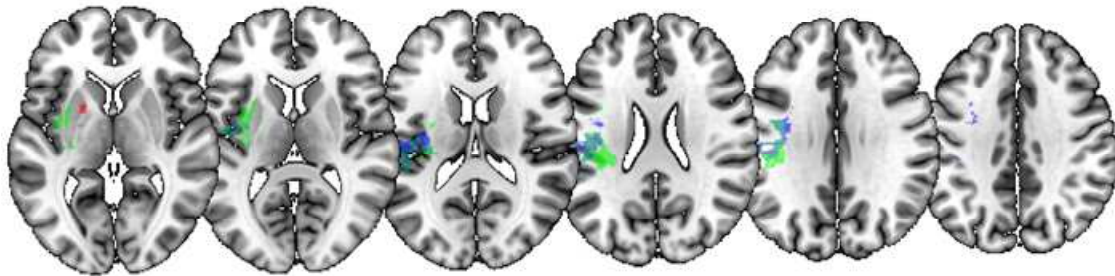


Figure 1.3. Results from the preliminary univariate VLSM analyses. Colors are as follows: Blue: distorted substitutions/additions; Green: increased intersegmental durations; Red: distorted substitutions

To inspect the role of structural network connectivity in AOS, an analysis with diffusion tensor imaging (DTI) data was completed for each of these three measures for 39 of the participants who had available DTI scans. Connections to the left precentral gyrus (PrCG) were significant in all DTI analyses, and regional connections between the

left IFGpo and the left PrCG were the most predictive of each of the three behaviors in question. These results are presented in Table 1.5 and Figure 1.4.

Table 1.5

Significant inter-regional connections obtained from preliminary DTI analyses*

	Distorted Additions/ Substitutions	Increased Intersegmental Durations	Distorted Substitutions
Significant Connections	L MFG – L IFGpt	L IFGpo – L IFGpt	L IFGpo – L PrCG
	L IFGpo – L IFGpt	L MFG – L PrCG	
	L MFG – L PrCG	L IFGpo – L PrCG	
	L IFGpo – L PrCG		
	L Front. Orb – L PrCG		
	L SFG – R PrCG		

*Note that directionality cannot be implied from this analysis

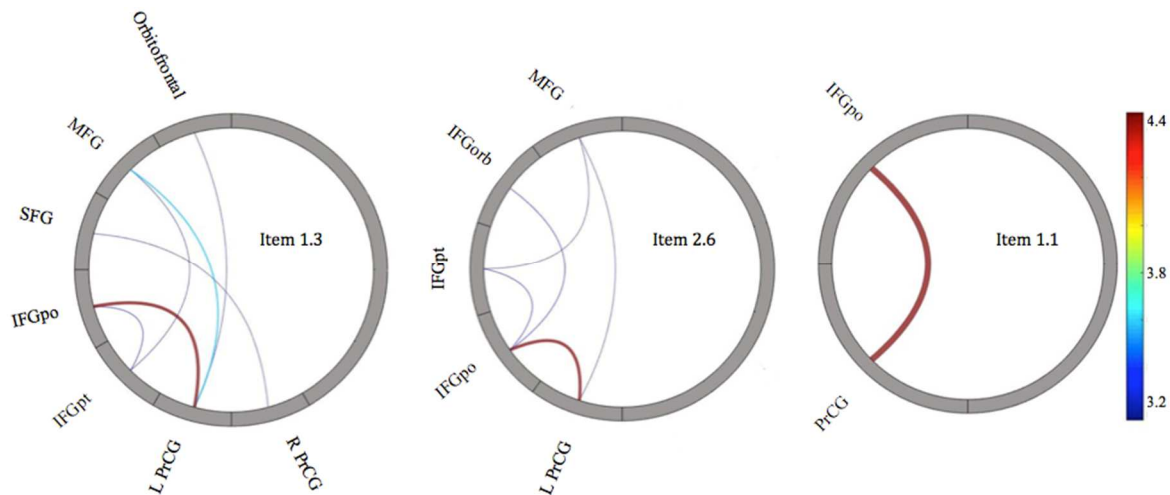


Figure 1.4. DTI results from the preliminary ASRS Items 1.3 (distorted substitutions/additions), 2.6 (increased intersegmental durations) and 1.1 (distorted substitutions). The left half of each circle graph corresponds to left hemisphere regions, the right half corresponds to right hemisphere regions. Warmer colors are indicative of greater connection strengths.

1.8.4 Summary. The behavioral results from these pilot data suggest that subjective measures (ASRS) have important explanatory value in classifying patients

with and without AOS. Results from the behavioral and imaging analyses suggest that individuals with speech production impairments are indeed a heterogeneous population, but that certain behavioral patterns emerge when extensive speech and language testing is obtained. Importantly, the pilot data show that patients can be distinguished by the nature of their impairments (i.e., phonetic vs. phonemic), and that individual behaviors can be dissociated neuroanatomically.

These preliminary analyses only included subjective measures (i.e., the ASRS). However, because clinicians may demonstrate low inter-rater reliability with AOS diagnosis (Haley et al., 2013; Haley et al., 2012), determining objective measures that classify individuals based on their presenting speech impairment can improve differential diagnosis of AOS, and allow for more accurate comparisons of results across studies. After all, McNeil and colleagues (1997) stated that due to potential differences in perceptual (subjective) diagnosis of AOS across clinicians and research groups, there is little reassurance that the findings of many early AOS studies have in fact investigated the same disorder. Furthermore, if objective acoustic measures can distinguish between patient groups, devising ways to implement and efficiently analyze acoustic measures in a clinical setting may provide practicing clinicians more valid and reliable ways to gather important speech production measures from speech samples.

To our knowledge, no study has used classification analyses along with neuroimaging data to determine behavioral measures that best distinguish groups of patients with post-stroke speech and language deficits, nor have these behaviors been related to patterns of brain damage. As such, the results from this study will contribute to the clinical and theoretical literature regarding speech production deficits in post-stroke

individuals. Clinically, there is a need for more reliable and valid testing measures to distinguish patients with AOS, aphasia, and dysarthria. Speech production behaviors that reliably distinguish clinical groups may be used to inform future studies aimed to create testing measures with greater sensitivity and specificity. Theoretically, relating the current findings to contemporary models of speech production may advance our understanding of the level(s) that speech planning/programming impairments occur. Finally, identifying patterns of damage to cortical regions and network connections that predict these speech behaviors will inform future studies investigating brain-behavior relationships that support speech production.

In the chapters that follow, experimental methods will be described thoroughly (Chapter 2), behavioral and neuroimaging results will be presented (Chapter 3), and finally, results will be summarized and discussed in the contexts of clinical practice and contemporary models of speech production (Chapter 4).

CHAPTER 2

METHOD

2.1 PARTICIPANTS

This study is retrospective; it includes data from individuals who have completed speech/language testing and neuroimaging within the Aphasia Lab at the University of South Carolina over the past 10 years. The study in which the current sample was obtained recruited individuals with a history of single-event ischemic stroke, in the chronic phase of recovery (i.e., six months post-stroke), and between the ages of 20-80 (both at time of stroke and testing). For the purpose of the current study, only individuals with left hemisphere stroke were considered. Exclusion criteria included history of neurological disorder affecting the brain (e.g., dementia), traumatic brain injury, or other speech-language impairment. In addition, individuals who had a diagnosis of severe aphasia (i.e., WAB AQ score <20) were excluded to ensure adequate speech production for transcription and scoring of speech and language errors.

From the initial sample of 77 individuals who completed testing, the final sample included here was composed of 58 participants (mean age = 60.53 ± 10.54 ; 21 females; months post-onset = 66.07 ± 52.93). 19 were excluded for the following reasons: aphasia severity (n=4), stroke not localized to left hemisphere cortical region (brainstem stroke: n=2; right hemisphere stroke: n=4; no evidence of stroke on MRI: n=1), concomitant neurological impairment (dementia: n=2; MRI consistent with anoxic event: n=1), history of developmental speech impairment (n=1), and four were excluded due to insufficient

data (unable to undergo neuroimaging: n=2; malfunction with video recording equipment: n=2). All participants included in this sample consented to study procedures by signing an informed consent form approved by the Institutional Review Board at the University of South Carolina. All participants were reimbursed for their time and compensated for travel.

Each participant was assigned to one of three groups determined by their ASRS and WAB AQ scores. Based on ASRS criteria, 20 participants demonstrated behaviors consistent with a diagnosis of AOS. Of these 20, all but two had a diagnosis of aphasia³ (according to WAB diagnostic criteria). These individuals were assigned the AOS-Aphasia group. The remaining 38 participants without AOS could be further classified with aphasia only (n=25 participants, the “Aphasia Only” group) or no chronic post-stroke communication impairment measured by WAB or ASRS criteria (n= 13; herein referred to as the “Stroke Control” group). Group characteristics are presented in Table 2.1 below, and independent samples t-tests indicate group differences where applicable.

³ The two individuals with AOS that did not test aphasic according to the WAB were included with the AOS-aphasia group.

Table 2.1

Characteristics of the study sample

	Stroke Control (n=13)	Aphasia Only (n=25)	AOS-Aphasia (n=20)	Significant Differences*
Mean Age	64.83±10.56	60.48±10.03	58.9±10.53	SC=AO=A-A
Male:Female	4:9	20:5	13:7	-
WAB AQ	98.43±1.12	75.9±20.36	64.19±21.41	SC>AO>A-A
Aphasia Types	N/A	Anomic: 11	Anomic: 4	-
		Broca's: 4	Broca's: 13	-
		Conduction: 7	Global: 1	-
		Wernicke's: 3	None: 2**	-
ASRS AOS Severity	0	0	2.8±0.95	SC=AO A-A>SC A-A>AO
ASRS Aphasia Severity	0	1.72±0.94	2.2±1.32	SC<AO SC<A-A AO=A-A
ASRS Dysarthria Severity	0	0.13±0.34	0.45±0.83	SC=AO SC<A-A AO=A-A

*Based on independent samples t-tests, $p<0.05$; SC: Stroke Control; AO: Aphasia Only;
A-A: AOS-Aphasia

**These two individuals were not aphasic according to WAB AQ cut-off criterion of 93.8

2.2 SPEECH AND LANGUAGE TESTING

As part of the study from which this sample was obtained, all participants completed an extensive battery of speech-language testing. A portion of the test battery included audiovisual recording of three connected speech samples and a motor speech evaluation. Details of these tasks and scoring procedures for this study are described below.

2.2.1 Connected Speech Samples. Three picture description tasks were used to measure speech production deficits during connected speech. The pictures used were black and white drawings depicting visually rich material, obtained from assessments commonly used in research and the clinical management of aphasia. The pictures were as follows: the “Cookie Theft” picture from the *BDAE* (Goodglass et al., 2000), the “Circus” picture from the *ABA-2* (Dabul, 2000), and the “Picnic Scene” from the *WAB* (Kertesz, 1982; Kertesz, 2007). Participants were allowed two-minutes per picture to describe the scene and were encouraged to use complete sentences and to include as much detail as possible.

2.2.2 Motor Speech Evaluation. The motor speech evaluation (Duffy, 2005) was used to evaluate various aspects of speech production to complete ASRS scoring for all participants. The following measures obtained from the motor speech evaluation were used to score specific ASRS items: a) word, phrase and sentence repetition tasks (used to score ASRS items pertaining to production at increasing utterance length and complexity; ASRS1.3, ASRS1.4), and b) measures of articulatory agility (alternating and sequential motor rates, i.e., repetition of “puh, puh, puh” and “puh tuh kuh”, respectively; used to score ASRS items 1.5 and 3.1). In addition to these production measures, an oral mechanism examination was used to determine if participants presented with concomitant

dysarthria. Movement of the articulators (lips, tongue, jaw) was evaluated using non-speech tasks to identify oral-motor weakness (associated with flaccid dysarthria) or spasticity (associated with upper motor neuron dysarthria). The motor speech evaluation is included in Appendix B.

2.2.3 Verbal Naming. For a subset of participants (n=38), the Philadelphia Naming Test (PNT) (Roach, Schwartz, Martin, Grewal, & Brecher, 1996) was used to obtain production measures at the single word level. The PNT is a 175-item evaluation of single-item picture naming, where participants are presented with a picture on a computer screen and are given 30 seconds to name the item. Only a subset of responses was scored, as described below (Section 2.4).

2.3 NEUROIMAGING DATA

MRI data were acquired using a Siemens 3T Trio System with a 12-channel head-coil. All participants underwent scanning with the following imaging sequences: 1. A T1-weighted imaging sequence using a MP-RAGE (TFE) sequence with a FOV=256x256mm, 192 sagittal slices, 9° flip angle, TR=2250ms, TI=925ms, and TE=4.15ms, GRAPPA=2, 80 reference lines; 2. A T2-weighted MRI for the purpose of lesion-demarcation with a 3D SPACE (Sampling Perfection with Application optimized Contrasts by using different flip angle Evolutions) protocol with the following parameters: FOV= 256x256mm, 160 sagittal slices, variable flip angle, TR=3200ms, TE=212ms, no slice acceleration. The same slice center and angulation was used as with the T1 sequence; and 3. A diffusion EPI scan (40 directions with s/mm^2 , TR=5000 ms, TE=79 ms, 82×82 matrix, 207×207 mm FOV, parallel imaging GRAPPA=2, 50 contiguous 2.3 mm axial slices, TA=390 s).

2.4 PROCEDURE

Audiovisual recordings from the aforementioned speech production tasks (picture descriptions, naming, motor speech evaluation) were used to obtain subjective and objective measures of speech production (described in detail below). Video recordings obtained from the three picture description tasks and the PNT were converted to WAV files (44 KHz sampling rate) and annotated using Praat sound analysis software (Boersma & Weenink, 2001). All annotations were completed by an American Speech-Language Hearing Association (ASHA)-certified speech-language pathologist (SLP) with experience in the assessment of individuals with post-stroke speech and language deficits. For the purpose of reliability, a subset of the data was scored by another ASHA-certified SLP with extensive experience with acoustic analysis and graduate research assistants trained in transcription and coding (see Section 2.4.8 for details regarding inter-rater reliability for all measures included in this study).

In addition to objective measures derived from annotations of recorded speech, subjective ratings of speech production were obtained using currently available measures (the ASRS, select portions of the WAB/WAB-R). Details of all measures are provided below.

2.4.1 Objective Measures. Pairwise variability index (PVI), voice onset time (VOT), amplitude envelope modulation spectrum (EMS), narrow transcription speech sound distortions (Cunningham et al., 2015) and measures of speech fluency (syllabic rate, words per minute; WPM; and speech productivity) served as objective measures of speech production. Each measure is described in detail below.

2.4.2 Pairwise Variability Index (PVI). PVI has been used as an objective measure of rhythm and timing of speech production in speakers with post-stroke AOS

(Ballard et al., 2016; Vergis et al., 2014) and primary progressive AOS (Duffy et al., 2015). This measure is obtained by computing the relative difference between vowel durations in multisyllabic words, as well as across phrases and sentences (Ling, Grabe, & Nolan, 2000). In AOS, atypical prosody is characterized by the production of multisyllabic words with reduced syllable contrastiveness, i.e., vowels are produced with “equal and excess” stress (McNeil et al., 1997). Vowel lengthening effects can also occur across word and phrase boundaries, meaning atypical prosody can manifest across different levels of speech production (i.e., words and phrases; Duffy, 2005; Strand et al., 2014; Vergis et al., 2014).

There are few objective measures to characterize speech rhythm in individuals with MSDs. In a small sample of patients with AOS and aphasia, aphasia only, and control individuals, Vergis et al. (2014) found that the PVI measures for vowel duration (henceforth PVI-V) were similar for the individuals with aphasia and the control group, but differed significantly for the group with aphasia *and* AOS. Specifically, the individuals with AOS and aphasia demonstrated reduced vowel duration contrasts for words that followed a weak-strong stress pattern (e.g., potato). This finding was replicated in a larger sample, showing that PVI-V may be instrumental in the classification of AOS (Ballard et al., 2016). Therefore, in the current study, PVI-V measures were obtained at the word-level (for a subset of participants, n=38) and for connected speech (all participants). To our knowledge, this measure has not been used to characterize connected speech in a large sample of post-stroke patients with and without AOS, as Ballard et al. (2016) and Vergis et al. (2014) investigated production at the word level.

In this study, PVI-V for single words and connected speech was used to confirm the utility of this measure in the differentiation of diagnostic class in a large, independent sample of post-stroke individuals. Further inspection of this measure was used to determine if a subset of participants with AOS can be grouped according to PVI-V, providing insight into the existence of AOS subtypes. That is, nPVI-V was used to determine if those with AOS could be distinguished by syllabic and phonetic errors (Duffy & Josephs, 2012; Duffy, Strand, & Josephs, 2014; Maas et al., 2015).

To obtain PVI coefficients, audio-recorded speech was annotated using Praat (Boersma & Weenink, 2001). All speech samples were segmented for C/V segments, regardless of number of syllables in the word, or segment position (Thomas & Carter, 2006). Segmentation was completed according to guidelines by Peterson and Lehiste (1960) with additional considerations for spontaneous speech (Thomas & Carter, 2006) and speakers with production disorders (Liss et al., 2009). All intervals were segmented based on visible formant structures (for vowel onset/offsets) and spectral energy corresponding to different consonant classes (for consonant onsets/offsets). The following sections detail specific criteria for vowel boundaries, consonant boundaries, and pauses.

Vowel Boundaries. Vowels were identified based on visible formant structure. Vowel onset boundaries were placed at the onset of visible voicing, corresponding to the onset of the second formant. Vowel offset boundaries were identified by termination of the formant structure or abrupt change in amplitude that preceded the onset of a consonant. Devoiced vowels, while infrequent, were not included in the vowel segments, but instead, with the adjoining consonant interval (Liss et al., 2009).

Consonant Boundaries. Consonant boundaries were identified by the onset of frication (fricative consonants), spectral energy that corresponded to burst release (plosives), nasal formant structure (nasals), and at the first formant for sonorant consonants (/l/, /r/, /w/). In cases where a sonorant preceded or followed a vowel, amplitude of the first formant was used to guide segmentation (Ordin & Polyanskaya, 2015). The end of each consonant boundary was based on termination of frication (fricatives), the first formant structure (sonorants), and for plosives in the final position, plosive release was captured in consonantal segments provided it was visible on the spectrogram (Ordin & Polyanskaya, 2015). Plosives that occurred in the medial position of a word, where a burst release was not visible (as in the production of a 'tap,' e.g., /bʌrə/ for *butter*), were identified by the segment of reduced spectral energy (White & Mattys, 2007). In cases where a plosive in the medial position could be identified by a stop gap, the period of silence preceding the plosive was not included in the plosive boundary; accordingly, if periods of silence preceded the onset of other consonantal segments, these silences were not included in the consonantal interval so as not to skew the duration of plosive segments (White & Mattys, 2007).

Pauses. Periods of silence were not included in C or V intervals. Pauses between words and phrases were not included in the calculation of PVI (Grabe & Low, 2002; Thomas & Carter, 2006; White & Mattys, 2007). Grabe and Low (2002) advocated for the omission of pauses from the calculation of the PVI coefficient to reduce bias that can be introduced when this measure is computed across word and phrase boundaries. Pertinent to the current study sample, utterance boundaries can be especially challenging to determine in speakers with aphasia, as pauses can occur due to linguistic and/or motor speech production difficulties. Therefore, in an effort to reduce subjectivity in the

segmentation process, C/V intervals were segmented only when visible on the spectrogram, according to the above stated criteria. Utterances in which extraneous background noise impeded accurate reading of the spectrogram were excluded from PVI calculation. An example Praat annotation for a speaker with AOS (without concomitant aphasia) is presented in Figure 2.1.

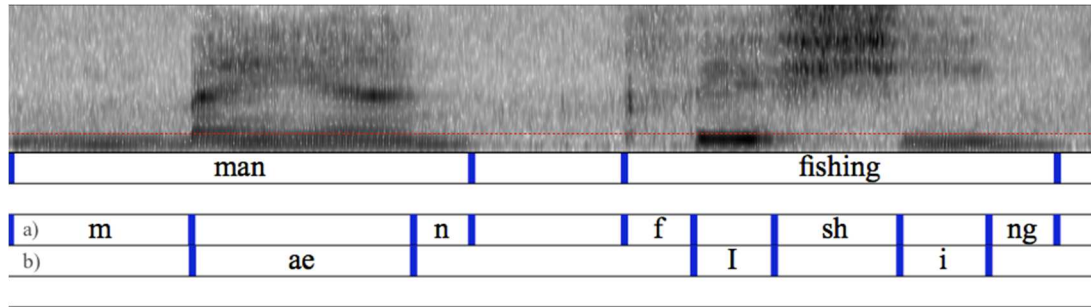


Figure 2.1. Example spectrogram from a speaker with AOS. The “Cons” and “Vowel” tiers are indicated by the “a” and “b,” respectively.

Calculation of PVI coefficients. PVI coefficients obtained from connected speech were rate normalized, herein referred to as nPVI-V. nPVI-V was calculated from the comparison of the duration of each successive vocalic interval using the following formula:

Equation 2.1

$$nPVI = 100 \times \left[\sum_{k=1}^{m-1} \frac{|d_k - d_{k+1}|}{|(d_k - d_{k+1})/2|} / (m - 1) \right]$$

Equation 2.1. Calculation of nPVI coefficient, where d=the duration of the selected interval, k is the chosen interval, and m=the interval number in which the PVI equation is calculated (Grabe & Low, 2002).

At the word level, raw PVI-V (rPVI-V) coefficients were calculated from the first two syllables of words that contained three or more syllables. rPVI was calculated for each word, and then averaged across all words with the same stress patterns (i.e., weak-

strong or strong-weak; Ballard et al., 2016; Vergis et al., 2004). Unlike the nPVI coefficients obtained for connected speech, the word-level rPVI coefficients were not rate normalized to facilitate comparisons with prior work (Ballard et al., 2016; Vergis et al., 2014).

Equation 2.2 was used to calculate rPVI-V values for the W-S and S-W words.

Equation 2.2

$$rPVI = 100 \times \left[\sum_{k=1}^{m-1} \frac{|d_k - d_{k+1}|}{m - 1} \right]$$

Equation 2.2. Calculation of rPVI coefficient. Variables are the same as those defined in Equation 2.1 (Grabe & Low, 2002).

2.4.3 Voice Onset Time. In plosive consonants (e.g., /b/, /d/, /g/), VOT is the temporal lag between the release of labial constriction and the onset of vocal fold vibration for voicing (Zlatin, 1974). VOT is considered a phonetic measure, as it is suggested to reflect speech motor control, specifically the coordination between glottal opening/closing, and movements of the lips, tongue, and jaw (Auzou et al., 2000; Baum, Blumstein, Naeser, & Palumbo, 1990; Shewan, Leeper, & Booth, 1984). Although several studies that have investigated VOT in aphasia and/or AOS exist, many of these studies include only single cases (Freeman, Sands, & Harris, 1978; Ziegler & von Cramon, 1985) or small sample sizes (Baum et al., 1990; Blumstein, Baker, & Goodglass, 1977; Blumstein, Cooper, Goodglass, Statlender, & Gottlieb, 1980; Freeman et al., 1978; Hoit-Dalgaard, Murry, & Kopp, 1983; Itoh et al., 1982; Kent & McNeil, 1987; Kent & Rosenbek, 1983; Odell, McNeil, Rosenbek, & Hunter, 1991; Seddoh et al., 1996; Shewan et al., 1984; Ziegler & von Cramon, 1985). In many of these studies, descriptions of the measures used to diagnose AOS and/or aphasia are not reported (for

review, see Wambaugh, Doyle, Kalinyak, & West, 1996). For these reasons, there is considerable variability in the results obtained from these studies, although some behavioral patterns have emerged. Generally, speakers with AOS demonstrate a lag in VOT for voiced plosives, meaning voicing occurs long after the release of the stop consonant, beyond what is average for typical speakers. Speakers with AOS may also demonstrate “voice lead,” voicing that precedes the release of the stop (Duffy, 2005). In either case, VOT tends to be longer in AOS than aphasia (Seddoh et al., 1996).

Previous research has shown that VOT in plosives is reliably different in speakers with AOS compared to those without AOS (Auzou et al., 2000; Baum et al., 1990; Blumstein, Alexander, Ryalls, Katz, & Dworetzky, 1987; Seddoh et al., 1996). Furthermore, VOT errors are affected by word position, with more errors occurring for word initial plosives (Freeman et al., 1978). Therefore, in the current study, VOT from word-initial plosives served as a measure of phonetic-level articulatory timing and coordination. Specifically, VOT variability served as the primary measure of interest as it has been suggested that individuals with AOS have greater VOT variability (Seddoh et al., 1995).

VOT was measured in Praat using spectrograms of participant audio recordings from the connected speech samples (all participants) and the PNT (n=38 participants). For both samples, VOT was obtained by manually demarcating the interval from the beginning of the burst release to the onset of voicing. In the connected speech samples, all word-initial plosives were measured. For the word-level analysis, a set of mono- and disyllabic words with word-initial plosives was obtained from the PNT. A complete list of target words obtained from the PNT is presented in Appendix C.

2.4.4. Amplitude Envelope Modulation Spectrum. Speech is characterized by low-rate modulations in amplitude that correspond to different aspects of the speech signal, ranging from rhythmic fluctuations associated with prosody and syllabic nuclei, to faster cycles associated with rapid articulatory movements (Crouzet & Ainsworth, 2001; Hall & Grose, 1993). These modulations can be characterized by the amplitude envelope modulation spectrum (EMS), where amplitude peaks that correspond to select frequencies can be used to describe dominant rhythms in an individual's speech (i.e., peak frequencies), and the degree that these rhythms dominate speech (i.e., peak amplitudes; Liss et al., 2010).

Much of the research pertaining to EMS has come from the speech perception literature (Crouzet & Ainsworth, 2001; Ghitza, 2011, 2013; Ghitza, Giraud, & Poeppel, 2012; Ghitza & Greenberg, 2009; Giraud et al., 2007; Giraud & Poeppel, 2012; Hall & Grose, 1993; Poeppel, 2003), with some preliminary findings regarding the utility of these measures to study disordered speech (Liss et al., 2010) and voice (Carbonell, Lester, Story, & Lotto, 2015) production. Collectively, the shape of the modulation spectrum may provide insight into the timing and coordination of speech production (Crouzet & Ainsworth, 2001; Ghitza, 2011, 2013; Liss et al., 2010). Slow rate modulations have been associated with speech rhythm - at syllabic and prosodic levels. Modulation at the rate of 4 Hz has been associated with the periodicity of syllabic production, corresponding to cycles of jaw opening and closure (Giraud & Poeppel, 2012; MacNeilage, 1998), and slower modulations, i.e., 1-3 Hz, have been associated with the prosodic contour of connected speech. Faster modulations are associated with phonetic features of speech. Particularly, Ghitza (2011) proposed that the beta range (15-

30 Hz) is associated with articulatory transitions, or dyad production (i.e., transition from one phone to the next), and that these rates are constrained by syllabic (~4 Hz) structure.

In the dysarthria literature, amplitude modulation, as characterized by several measures (e.g., peak frequency, peak amplitude, and energy in the 3-6 Hz, 0-4 Hz, and 4-10 Hz spectra; Liss et al., 2010), has been shown to be a reliable predictor in distinguishing speakers with dysarthria from normal controls as well classifying individuals into the different dysarthria subtypes (Liss et al., 2010). Moreover, it has been suggested that EMS measures may be clinically useful. In the dysarthrias, some of these metrics are related to PVI (which also demonstrates utility dysarthria classification), but unlike PVI, can be obtained from speech samples without manual segmentation of speech (Liss et al., 2010). To our knowledge, no published studies have investigated the utility of EMS in the differential diagnosis of AOS. Therefore, EMS measures were included here to characterize apraxic and aphasic speech, particularly at syllabic, prosodic, and articulatory levels.

To this end, EMS measures were obtained from the three connected speech samples. All audio recordings were preprocessed using Adobe Soundbooth to remove noise and other extraneous sounds (e.g., clinician interjection, background noise). Subsequent procedures were carried out in Matlab using custom scripts. EMS was calculated from the first 90 seconds of each sample, over three, 20 sec windows at the beginning (from 0-20 sec) middle (35-75 sec), and end (70-90 sec) of each sample. The purpose of the three windows was to observe variability in EMS measures over the course of each speech sample. As well, only the first 90 seconds was used from each sample, as there was a trend for participants to decrease speech production towards the end of the two-minute speaking duration.

For each 20-second sample, speech was downsampled to a frequency of 16 kHz. Envelopes were extracted via halfwave rectification, and the envelope was low-pass filtered using a 6th order butterworth filter at 50 Hz. The low-pass filtered envelope was then downsampled to a sampling frequency of 1000 Hz. Next, the Fast Fourier transform (FFT) was computed. Energy was summed into FFT bins that corresponded to octave bands with the center frequencies ranging from 1-32 Hz. Finally, the energy in an octave-band FFT bin was divided by the energy in the 0 bin, yielding the modulation index relative to the DC offset. A number of dependent variables were derived from the individual octave bands or the full signal, some in replication of Liss et al. (2010). An example EMS plot for a speaker with AOS and aphasia is presented in Figure 2.2, and Table 2.2 provides an explanation of each EMS measure inspected in this study.

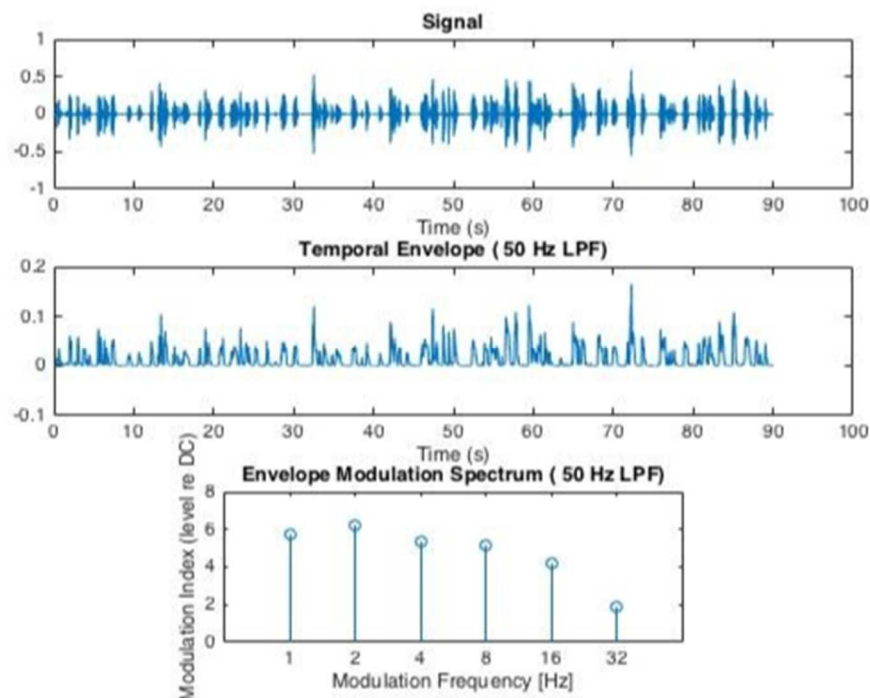


Figure 2.2. Example EMS plot for a speaker with AOS. The top portion of each panel depicts the raw audio file, the middle portion depicts the extracted temporal envelope, and the bottom portion displays the modulation index for each frequency band.

Table 2.2

EMS variables used in analyses

Measure	Description
Peak Frequency	The frequency corresponding to the spectral peak across the 1-32 Hz range. This reflects the dominant rate of each participant's amplitude pattern (Liss et al., 2010). In typical speakers, peak frequency tends to fall between 3-6 Hz, corresponding to the rate of syllable production (Giraud et al., 2007)
Peak Amplitude	The amplitude of the peak frequency, representing the degree to which the rhythm was dominated by this frequency
1Hz Band 2 Hz Band	Amplitude in the a) 1 Hz and b) 2 Hz octave bands, each divided by the total energy of the modulation spectrum. Amplitude fluctuations at lower frequencies reflect the prosodic features of an utterance*
4Hz Band 8 Hz Band	Amplitude in the a) 4 Hz and b) 8 Hz octave bands, each divided by the total energy of the modulation spectrum. These frequencies are suggested to correspond to the rate of syllabic production*
16 Hz Band 32 Hz Band	Mean amplitude of the a) 16 Hz and b) 32 Hz bands, each divided by the total energy of the modulation spectrum. These frequencies are arguably reflective of the duration of phonetic features that occur in speech production*

*Based on work by Giraud et al., 2007; Ghitza, 2011

2.4.5 Narrow Phonetic Transcription. Phonetic errors characterized by distortions are commonly attributed to AOS, but few studies have explicitly evaluated the rate and frequency of different types of distortion errors in AOS compared to speakers without AOS. A recent study by Cunningham and colleagues (2015) used narrow phonetic transcription in speech samples obtained from patients loosely defined as “possible-AOS” (P-AOS, based on measures of syllable segmentation) and “possible-phonemic paraphasia” (P-APP, based on perceptual ratings of phonemic production errors). Cunningham et al. (2015) found that the most common errors that distinguished

these groups included voicing errors, segmental lengthening errors, and tongue placement errors (e.g., dentalized, backing/fronting, palatized, frictionalized). Across these errors, the P-AOS group demonstrated greater error rates when compared to the P-APP group. Accordingly, distorted productions resulting from incorrect tongue placement (as reported by Cunningham et al., 2015) were identified and narrow transcription of these distortions was completed as part of the Praat annotations for the connected speech samples⁴. A list of narrow transcription codes can be found in Appendix D. Additionally, other sound level errors, i.e., phonemic paraphasias, were coded to obtain a measure of non-distorted, but incorrect, segmental production errors.

2.4.6 Measures of Speech Fluency. Lastly, total words per minute (WPM), syllabic rate, and speech productivity (Park et al., 2012) were obtained from the connected speech samples as objective measures of overall speech fluency. WPM was calculated by dividing the total number of words produced during the speech sample by the duration of the sample (i.e., two minutes). Syllabic rate was obtained by dividing the number of syllables by the total speaking time (in seconds) for each speech sample. Speech productivity was derived by dividing the total duration of the speech sample by the time spent producing speech, i.e., total duration of speech production, minus pauses (Park et al., 2011). Many measures exist to quantify speech fluency, and there is no single measure generally accepted to represent speech fluency. Therefore, these measures were chosen because they have been shown to be predictive of clinician classification of speech fluency (Park et al., 2011), or are frequently discussed with regard to variables obtained from the EMS (i.e., syllabic rate). As well, WPM has been shown to correlate

⁴ Voicing and segmental lengthening errors were not coded with narrow transcription, as these behaviors were captured via VOT and nPVI analysis.

with the WAB fluency rating scale (Basilakos et al., 2014), a measure widely used in clinical practice.

2.4.7 Perceptual Measures. The ASRS (Strand et al., 2014) was completed for each participant to classify speech production errors that occur exclusively in AOS, AOS and aphasia, AOS and dysarthria, and all three disorders. All participants were rated on the presence/severity of all speech characteristics on the ASRS based on a 5-point scale (0=not present; 1=detectable but not frequent; 2=frequent but not pervasive; 3=nearly always evident but not marked in severity; 4=nearly always evident and marked in severity). The ASRS scale can be found in Appendix A, and a summary of ASRS scores for all three participant groups is presented in Table 2.1. Aside from clinical classification, the purpose for including the ASRS measures was to determine the extent that objective measures correlate with perceptual ratings, and importantly, to identify the perceptual correlates of the objective measures that carry the greatest discriminative weight in discriminant analysis. Table 2.3 provides a summary of all study measures.

Table 2.3

Measures used to evaluate phonemic, phonetic, and prosodic impairments

	Phonemic	Phonetic	Prosodic	Other
Perceptual	-ASRS Items 3.1, 3.2 -Phonemic errors transcribed from speech sample	-ASRS Items 1.1-1.6	-ASRS Items 2.1, 2.2, 2.6	--
Objective	--	-VOT (PNT and picture description) -16 Hz band prominence	-PVI (PNT and picture description) -1 Hz/4 Hz band prominence	-WPM -Syllabic rate -Speech productivity

2.4.8 Reliability. Inter-rater reliability for measures obtained from the connected speech samples was established using a two-way mixed consistency single-measures intraclass correlation coefficient (ICC).⁵ Speech samples from six individuals (10.3% of the study sample) were randomly selected and the secondary rater was blind to participant characteristics that could influence ratings (i.e., diagnosis of aphasia or AOS). ICC values for each measure are as follows: n-PVI: 0.83; distortion errors: 0.87; phonemic errors: 0.88; and mean VOT (collapsed across voiced and voiceless targets): 0.98. All Chronbach's alpha values were >0.90. These ICC values are considered "excellent" (Cicchetti, 1994) and are in line with reliability measures reported in other AOS studies (e.g., Cunningham et al., 2015; Vergis et al., 2014). Inter-rater reliability for the ASRS scale has been reported previously (Basilakos et al., 2015: ICC=0.88; Moser et al., 2016: ICC=0.94).

Word-level measures. For the word level analyses completed by the trained master's students, the primary rater for the connected speech samples measured C/V segments and VOT for all targets from six randomly selected PNT assessments (17.1% of the study sample that were given the PNT). ICC values are as follows: PVI-V: 0.93; VOT (collapsed across voiced/voiceless initial consonants): 0.98 (Chronbach's alpha > 0.95 for both analyses).

2.5 DATA ANALYSES

2.5.1 Specific Aim 1: Behavioral Analyses. Group differences in behavioral measures were analyzed using SPSS version 22. Descriptive statistics (means and

⁵ Inter-rater reliability was calculated for all variables in their "final form," as advised by Hallgren (2012).

standard deviations) for all objective and subjective measures were computed, and analyses of variance (ANOVAs) were conducted between the AOS-Aphasia group, the Aphasia Only group, and the Stroke Control group to determine if group differences existed for the measures investigated. For variables that violated assumptions of normality (determined from Shapiro-Wilk p -values and visual inspection of distributions), non-parametric Kruskal-Wallis ANOVAs were conducted. All analyses were Bonferroni corrected for multiple comparisons by dividing the total number of variables by the standard p value of 0.05. Bonferroni corrected p -values for each analysis are presented in the Results sections pertaining to each analysis (Chapter 3).

Predicting Group Membership. Subsequent classification analyses were completed using Fisher's linear discriminant analysis (LDA). The rationale for this approach is that it can be used to determine if a given set of variables can predict group assignment, and if so, the measures that have the greatest classification weight can be determined. Essentially, this analysis also provided a metric with which to compare the discriminative weight of each input variable. Because a large number of variables was obtained in this study, many of which are significantly correlated (see Results Section 3.2 and correlation tables presented in Appendix G), not all variables were entered into the LDA model to avoid issues pertaining to statistical power. Therefore, variables entered into the LDA were chosen based on 1) theoretical rationale, 2) correlation with other speech production measures, and 3) correlation with aphasia severity (WAB AQ scores). Additional details regarding predictor selection can be found in the Results section (Section 3.3).

All LDA analyses were completed in Matlab using the Statistics and Machine Learning toolbox and custom scripts. Data were scaled between the values of 0-1 by dividing each score by the maximum possible score that could be obtained for that measure. For measures where this was not possible (i.e., nPVI-V, peak frequency), data were scaled first by calculating the absolute difference between each participant's value and a reference value obtained from the literature, and second, by dividing each absolute difference score by the largest difference score for that measure. As such, a reference score of 66 was used for the nPVI-V coefficient (Arvaniti, 2012), and 4 Hz was used as the reference value for peak frequency. Finally, variables were flipped (when necessary) so that a higher value corresponded to better performance. Variables in the opposite direction (i.e., where higher values corresponded to worse performance) were reversed by subtracting the score from 1.

Once all data were appropriately scaled, behavioral analyses were completed as follows:

First, preliminary analyses were completed with a multidimensional scaling (MDS) approach. All behavioral measures were entered for each participant (henceforth considered the independent variables), along with a binary classification of AOS or no AOS (the dependent variables; based on ASRS scores). MDS is used for visualizing distances between given observations. Here, it was used to visualize distances between the AOS or no AOS groups based on speech production measures.

To visualize the data with MDS, participant data were entered into an $n \times m$ dimension matrix, where n =number of participants and m =number of measures entered. Coordinates for these data points were then obtained based on straight-line (Euclidean)

distances, forming a distance matrix. Further decomposition of the distance matrix yields a new matrix that contains a set of coordinates used to plot the data based on obtained distances. Essentially, this step determined if participants could be distinguished into linearly distinct groups. Based on the pilot data reported in Chapter 1, it was predicted that groups would be linearly separated, meaning AOS can indeed be distinguished behaviorally from aphasia. As such, the MDS results should indicate shorter distances between individuals within the suspected AOS group, shorter distances between those without AOS, and greater distances between these groups.

Next, Fisher's LDA served as a means of hypothesis testing for Hypotheses 1a and 1b – the extent that obtained measures significantly explain group differences (1a) and whether prosodic measures indeed improve classification of group accuracy (1b). Here, failure to reject the null hypothesis pertaining to 1a would suggest that the predictor variables did not yield a linear, binary classification. In contrast, rejecting the null hypothesis would suggest that performance on the predictor variables indeed yielded two distinct groups. Importantly, the LDA was used to provide further detail regarding the extent that the subjective and acoustic measures obtained in this study (the independent variables) distinguished group membership, (i.e., AOS versus no AOS, dependent variables) as well as the independent variables that account for the greatest amount of variance between groups.

With regard to Hypothesis 1c specifically, measures of prosody (i.e., PVI, 1 Hz band prominence) were scrutinized to determine if they accounted for variance within the AOS group specifically. The purpose of this was to determine if subtypes of AOS do indeed exist, and if so, the LDA analysis with prosodic and phonetic errors (see Table

2.3) as independent variables was used to determine the extent these variables predicted groupings within the suspected AOS sample (i.e., AOS group 1 and AOS group 2; see Section 3.3 for further details).

Importantly, the LDA implemented a leave-one-out cross-validation (LOOCV) approach to avoid overfitting the model. This is an iterative process where x linear regression analyses were conducted, where x =the sample size, N . In each analysis, the data from $N-1$ participants was used to predict the classification of the n^{th} participant. Average prediction accuracy was obtained following the N runs, indicating how often group classification was accurately predicted from the LDA model.

2.5.2 Specific Aim 2: Neuroimaging Analyses. All participant images were preprocessed using the anatomical template in the Clinical Toolbox (Rorden, Bonilha, Fridriksson, Bender, & Karnath, 2012) for SPM12 (Ashburner et al., 2012). Lesions were demarcated by a neurologist in MRICron on individual T2-MRIs (in native space), using the T1-MRI and diffusion sequences for guidance. Preprocessing began with the co-registration of the T2-MRI to match the T1-MRI, aligning the lesions to native T1 space. Next, image segmentation and normalization was completed with enantiomorphic normalization (Nachev, Coulthard, Jäger, Kennard, & Husain, 2008) in SPM12 and custom MATLAB scripts.

Analyses. All subsequent neuroimaging analyses were completed using in-house Matlab routines (nii_stat software; <http://www.nitrc.org/projects/niistat/>). The neuroimaging analyses were completed in two steps. First, to identify patterns of cortical damage predictive of these behaviors, VLSM analyses were conducted for each variable in question. Aphasia severity (which is also highly correlated with lesion size; Wu et al.,

2015) was entered as a nuisance regressor to identify areas predictive of the speech behaviors, while controlling for aphasia severity. Note that this was done for all variables; regardless of the correlation (or lack thereof) with aphasia severity. Familywise error rate was controlled using permutation thresholding (2000 permutations; Winkler, Ridgway, Webster, Smith, & Nichols, 2014). For variables that did not survive thresholding when WAB AQ scores were included, a follow-up univariate analysis was completed (also corrected with 2000 permutations).

Next, a connectome analysis was used to a) obtain measures of regional white matter integrity that predict production impairments, and b) determine the extent that compromised network connectivity differs between participant groups (AOS versus those with aphasia only, or those without any detectable speech and/or language impairment). The connectome analysis uses information obtained from diffusion tensor imaging (DTI). DTI images are used to measure the magnitude of water diffusion in the brain, providing an estimate of the location of white matter pathways (Hagmann et al., 2010; Sporns, Tononi, & Kötter, 2005). A matrix of whole brain white matter connections, where each region can be connected to all other regions, is constructed from these measurements. In the current study, average connectome maps were constructed to measure network integrity across the brain, and relate white matter integrity to aspects of impaired speech production (Bonilha, Nesland, et al., 2014).

Whole brain connectome maps were constructed for all participants from structural MR and diffusion images. This process has been detailed in Bonilha et al. (2014), and used the following procedures: 1) Grey and white matter tissue maps generated via normalization routines completed with the Clinical Toolbox for SPM (Rorden et al., 2012) are used for creation of cortical ROIs and white matter maps. 2)

Grey matter maps are parcellated into ROIs based on co-registration to the JHU Atlas (Faria et al., 2012). Regions located within the drawn lesion boundaries are masked to avoid the parcellation of damaged tissue into ROIs. 3) Structural images (T1, T2) are co-registered to the B0 diffusion image, with the same transformations applied to all subsequent DTI volumes. This step ensures that grey and white matter maps are aligned to diffusion space, obtained from the DTI scans. 4) DTI images are reconstructed, using FSL (Smith et al., 2004), guided by the segmented white matter maps. 5) The number of fibers connecting each set of JHU ROIs is calculated to obtain a measure of total fiber counts between ROIs, and used to construct each individual connectivity matrix.

Because a great deal of inter-subject variability in whole brain white matter connectivity is likely, an exploratory analysis based on regions and connections implicated in contemporary models (i.e., HSFC, DIVA) was conducted. Here, fiber count connections between regions of interest implicated in speech production, along with white matter pathways previously shown to predict production impairments (Basilakos et al., 2014; Catani et al., 2012; Catani & Mesulam, 2008; Catani et al., 2013; Fridriksson, Guo, Fillmore, Holland, & Rorden, 2013), were included in correlation analyses with each behavioral variable. Additional details can be found in Section 3.4.

CHAPTER 3

RESULTS

3.1 BEHAVIORAL DATA: GROUP PERFORMANCE ACROSS MEASURES

Prior to statistical analyses, all variables were inspected for normality. At least one group demonstrated a distribution that did not meet the assumption of normality (determined by visual inspection of data and Shapiro Wilk $p < 0.05$) for the following variables (with group not meeting assumption in parentheses): variability in VOT for voiced stop consonants (VOT-SD_{voiced}; all groups), word-level rPVI and VOT-SD_{voiced/voiceless} (Aphasia Only, AOS-Aphasia), the proportions of phonemic and distortion errors to total words produced (all groups), syllabic rate (Stroke Control group only), mean WPM (Stroke Control group), and three of the measures derived from the amplitude modulation spectra (1 Hz band prominence: Aphasia Only group; 4 Hz band prominence: Stroke Control group; peak frequency and peak amplitude: all groups). No group distributions violated the assumption of normality for SD-VOT_{voiceless} and band prominence for 2 Hz, 8 Hz, 16 Hz, and 32 Hz. Statistical analyses for each variable are presented in the sections that follow.

3.2.1 PVI in Connected Speech. Mean nPVI-V coefficients are displayed in Figure 3.1. Means and standard deviations are as follows: Stroke Control: 62 ± 3.44 ; Aphasia Only: 61.42 ± 5.57 ; and AOS-Aphasia: 52.21 ± 7.10 . Notably, the Stroke Control and Aphasia Only group's nPVI-V coefficients are in line with values that have previously been reported for typical English speakers during connected speech. That is,

Arvaniti (2012) reported a mean nPVI-V of 66 for a group of English speaking adults during spontaneous speech.

A one-way ANOVA was conducted to determine if nPVI-V coefficients differed between the three groups. Results showed a significant effect of group, $F(2, 57)=18.11$, $p<0.001$, partial $\eta^2=0.40$. Post-hoc, Bonferroni corrected pairwise comparisons ($p=0.05/3$ groups= 0.0167) indicate that the AOS-Aphasia group had significantly smaller nPVI-V coefficients compared to the Stroke Control group [$t(31)=5.02$, $p<0.001$] and the Aphasia Only group [$t(43)=4.68$, $p<0.001$]. There was no significant difference in nPVI-V between the Stroke Control and Aphasia Only groups [$t(36)=1.21$, $p=n.s.$].

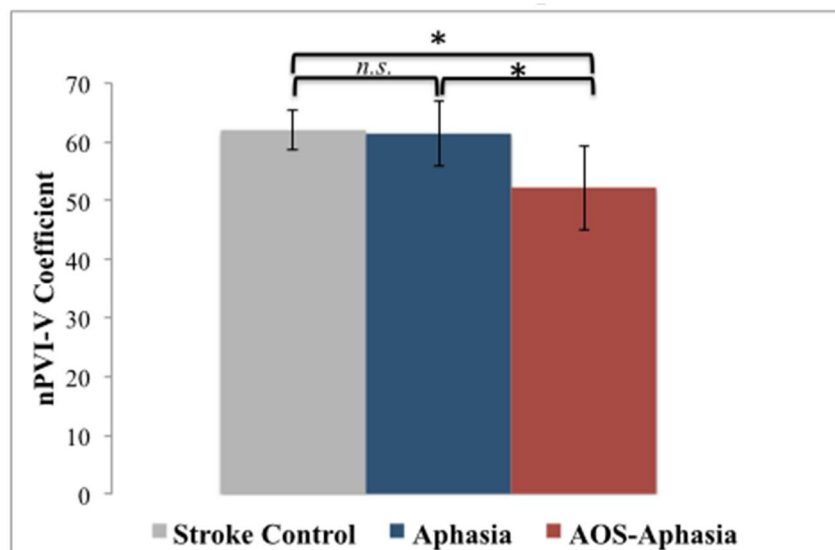


Figure 3.1. Mean nPVI-V coefficient values for each group. The asterisks indicate significant between-groups differences.

Because it has been suggested that individuals with AOS may be differentiated by distortion errors or by atypical prosody, the nPVI-V coefficients for the AOS-Aphasia group were inspected to determine the extent that nPVI-V was reduced similarly across participants. To this end, the AOS-Aphasia group was divided into those who had nPVI-V coefficients *within* one standard deviation of the mean of the Aphasia Only group

($66.53 > \text{nPVI-V} < 55.64$), below one standard deviation ($50.23 > \text{nPVI-V} < 55.82$), and below two standard deviations ($\text{nPVI-V} < 50.23$). Four participants that had nPVI-V coefficients within one standard deviation of the mean of the Aphasia Only group (mean= 57.78 ± 3.47), eight individuals had nPVI-V coefficients below one standard deviation (mean= 52.56 ± 1.76), and seven individuals were at least two standard deviations below the mean of the Aphasia Only group (mean= 45.71 ± 3.32). It should be noted that one participant in the AOS-Aphasia group had an nPVI-V score above one standard deviation of the Aphasia Only group. Figure 3.2 presents means for each of these “subgroups” compared to the Aphasia Only group. Further analysis regarding the extent that measures obtained from this study predict AOS “subgroups” is discussed in detail in Section 3.3.3.

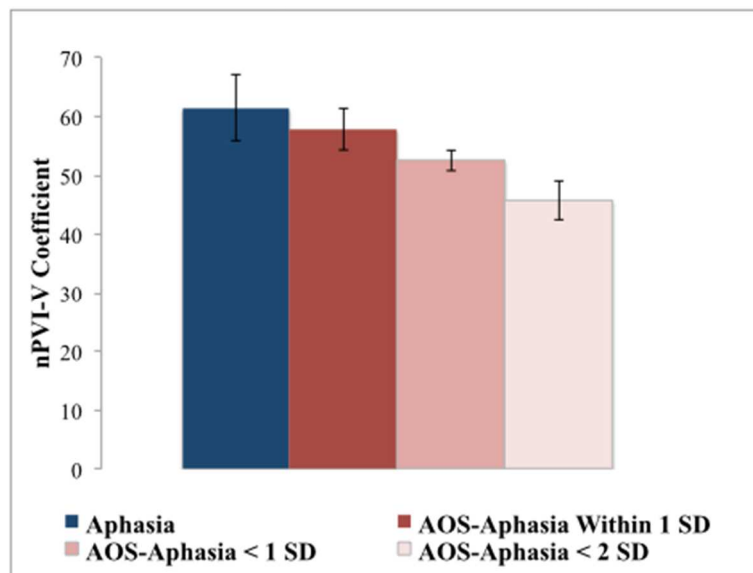


Figure 3.2. Mean nPVI-V coefficients for each AOS “subgroup.” Subgroups were derived from nPVI-V scores relative to the mean and standard deviation of the Aphasia Only group’s nPVI-V scores. Subgroup sample sizes are as follows: Within 1 SD: n=4; Below 1 SD: n=8; Below 2 SD: n=7.

3.2.2 Word-Level PVI. Due to differences in naming abilities, the number of items used to obtain word-level rPVI coefficients varied among participants (mean

targets produced=20 out of 21 possible, range=1-42, as multiple attempts were scored).

Notably, Ballard et al. (2016) reported that rPVI coefficients obtained from their participants could have been derived from as few as one spoken word.⁶

Mean rPVI coefficients for word level production are as follows: PVI-WS: Aphasia Only group: 55.15 ± 30.25 ; AOS-Aphasia group: 48.51 ± 20.95 . PVI-SW: Aphasia Only group: 47.49 ± 15.7 ; AOS-Aphasia: 39.36 ± 16.62 . There were no statistically significant for differences in mean rPVI coefficients at the word level (rPVI-WS: $U=147$, $z=0.98$, $p=n.s.$; rPVI-SW: $U=126$, $z=1.59$, $p=n.s.$). As indicated in Figure 3.3, rPVI coefficients were highly variable, particularly for the Aphasia Only group's PVI-WS coefficients. However, there were no significant group differences in rPVI variability (rPVI-WS: $U=167$, $z=0.40$, $p=n.s.$; rPVI-SW: $U=103$, $z=1.65$, $p=n.s.$) or in the range of rPVI coefficients for either word stress type, rPVI-WS: Moses span=34, $p=0.17$; rPVI-SW: Moses span=35, $p=0.73$.

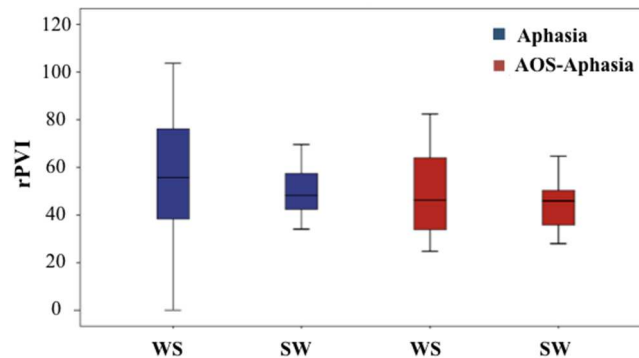


Figure 3.3. Box plots for rPVI-WS and rPVI-SW for both groups.

⁶ Ballard et al. (2016) state that some participant's rPVI coefficients were derived from as few as one token. Their rationale for including those with so few productions was based on prior work suggesting that this could be sufficient in the detection of group differences in PPAOS and typical speakers (e.g., Duffy et al., 2015).

Despite the lack of group differences, word-level rPVI was correlated with ASRS AOS severity, but only when rPVI was collapsed across stress type ($r_s = -0.72$, $p < 0.001$; corrected for aphasia severity: $r_s = -0.46$, $p < 0.01$). rPVI-WS was correlated with nPVI-V obtained from the connected speech samples ($r_s = 0.54$, $p < 0.001$ when controlling for aphasia severity); rPVI-SW was not ($p > 0.05$).

3.2.3 VOT in Connected Speech. The variability in VOT for voiced and voiceless initial stop consonants (VOT-SD_{voiced} and VOT-SD_{voiceless}, respectively) was analyzed to determine if individuals with AOS indeed demonstrate greater variability in articulatory timing and coordination. Mean variability \pm standard deviations are as follows: VOT-SD_{voiced}: Stroke Control = 0.014 ± 0.006 , Aphasia Only = 0.020 ± 0.01 , AOS-Aphasia = 0.033 ± 0.03 ; VOT-SD_{voiceless}: Stroke Control = 0.024 ± 0.01 ; Aphasia Only = 0.029 ± 0.01 ; AOS-Aphasia = 0.033 ± 0.014 . The distribution for VOT-SD_{voiced} violated the assumption of normality (Shapiro-Wilk $p < 0.05$); however, the distribution for VOT-SD_{voiceless} did not. Therefore, a Kruskal-Wallis one-way ANOVA was used to analyze the VOT-SD_{voiced} variables, and a parametric one-way ANOVA was used for the VOT-SD_{voiceless} comparisons.

Results of the Kruskal-Wallis one-way ANOVA of VOT-SD_{voiced} showed a significant effect of group, $X^2(2) = 8.06$, $p = 0.02$. Pairwise comparisons with Mann-Whitney U tests revealed a significant difference between the AOS-Aphasia group when compared to the Stroke Control group (AOS-Aphasia mean rank = 19.75; Stroke Control mean rank = 11.08; $U = 55$, $z = -2.53$, $p < 0.01$), but not when compared to the Aphasia Only group (AOS-Aphasia mean rank = 25.83; Aphasia Only mean rank = 18.67.28; $U = 153$,

$z=-1.86, p=0.06$). There was no significant difference between $VOT-SD_{voiced}$ for the Stroke Control and Aphasia Only groups ($p=0.11$).

Results of the one-way parametric ANOVA for $VOT-SD_{voiceless}$ were marginally significant [$F(2, 56)=3.26, p<0.05$], driven by a significant difference between the Stroke Control and AOS-Aphasia groups [$t(31)=2.56, p=0.016$]. There were no other statistically significant comparisons. The results for VOT variability for voiced and voiceless stop consonants are presented in Figure 3.4 below.

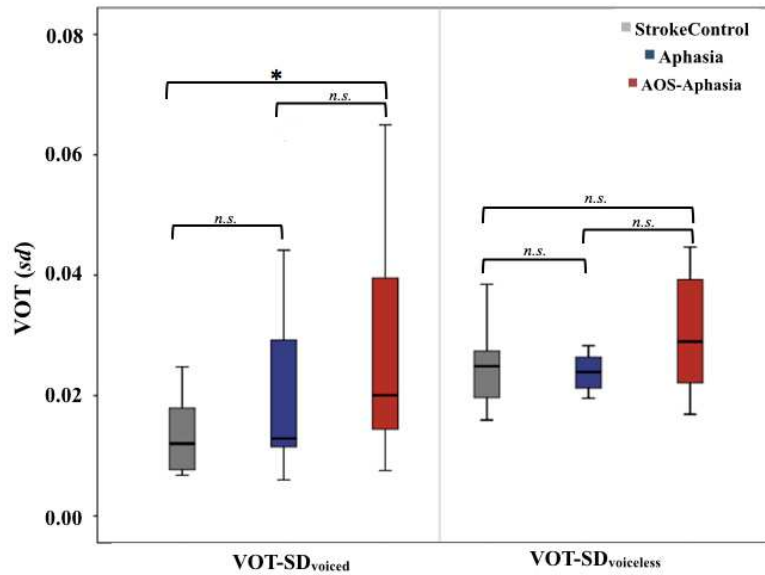


Figure 3.4. VOT variability measured from the connected speech samples. Significant group differences are indicated by asterisks.

3.2.4 Word-Level VOT. VOT from word-level naming attempts was obtained from a selection of 46 words (23 with initial voiced and 23 with initial voiceless stop consonants) that most participants generally produced, or attempted, on the PNT (See Appendix C for a complete list of target words). As with the word-level PVI coefficients, not all individuals produced every possible item, and in the case of multiple attempts at target production, all valid attempts were included. Mean targets produced are as follows:

voiced targets: 27.92 (range 9-58), voiceless targets: 17.14 (range 6-36). Mean variability \pm standard deviations for word-level VOT are as follows: VOT-SD_{voiced}: Aphasia Only=0.04 \pm 0.02, AOS-Aphasia=0.05 \pm 0.03; VOT-SD_{voiceless}: Stroke Control=0.03 \pm 0.01; Aphasia Only=0.029 \pm 0.01; AOS-Aphasia=0.031 \pm 0.01.

Figure 3.5 shows box plots for the variability in voiced and voiceless stop consonants for both groups. Although this figure indicates that the AOS-Aphasia group had a large range in VOT variability for voiced stop consonants, results from the Kruskal-Wallis ANOVA showed that there were no significant differences in VOT variability for either consonant type (i.e., voiced or voiceless; $p>0.05$). Furthermore, analysis of the ranges of VOT variability did not reveal any significant group differences (Moses span, $p>0.05$ for all comparisons) for voiced or voiceless stop consonants.

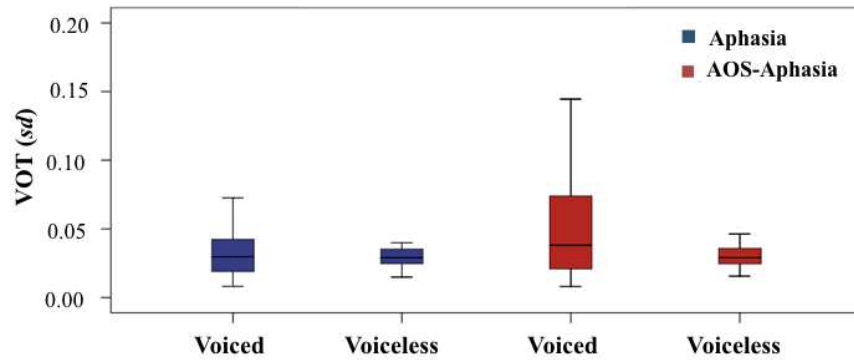


Figure 3.5. VOT variability measured from word-level production. There were no significant between-groups differences.

3.2.5 Phonemic and Distortion Errors. Phonemic and distortion errors were analyzed as the proportion of errors (phonemic or distortion) per total words produced. The proportions of phonemic errors and distortion errors are as follows: Stroke Control: phonemic errors=0.002 \pm 0.003, distortion errors=0.005 \pm 0.008; Aphasia Only: phonemic errors=0.03 \pm 0.04, distortion errors=0.02 \pm 0.03; and AOS-Aphasia: phonemic errors=

0.11±0.09, distortion errors=0.19±0.17. Between groups comparisons were analyzed using independent-samples Kruskal-Wallis ANOVAs. The distributions of error proportions differed significantly across groups, both for phonemic errors, $\chi^2(2)=32.48$, $p<0.001$, and distortion errors, $\chi^2(2)=34.94$, $p<0.001$. Results are presented in Figure 3.6.

Post-hoc pairwise comparisons with Mann-Whitney U tests (Bonferroni corrected level of significance= $p<0.0167$) show significantly higher rates of phonemic errors for the AOS-Aphasia group compared to both the Aphasia Only (Aphasia Only mean rank=16.80 vs. AOS-Aphasia mean rank=30.75, $U=95.5$, $z=-3.5$, $p<0.001$) and Stroke Control groups (Stroke Control mean rank=7.77 vs. AOS-Aphasia mean rank=23, $U=10$, $z=-4.44$, $p<0.001$). The Aphasia Only group also produced significantly more phonemic errors than the Stroke Control group (Aphasia Only mean rank=25.36; Stroke Control mean rank=8.23, $U=16$, $z=-4.51$, $p<0.001$).

The AOS-Aphasia group had significantly more distortion errors when compared to the other two groups. Results of the pairwise comparisons with the AOS-Aphasia group are as follows: Stroke Control (mean rank=7.15) vs. AOS-Aphasia (mean rank=23.4), $U=2$, $z=-4.72$, $p<0.001$. Aphasia Only (mean rank=16.8) vs. AOS-Aphasia (mean rank=34), $U=30$, $z=-5.03$, $p<0.001$. There were no significant differences in distortion errors between the Stroke Control and Aphasia Only groups, $U=95$, $z=-2.09$, $p>0.0167$.

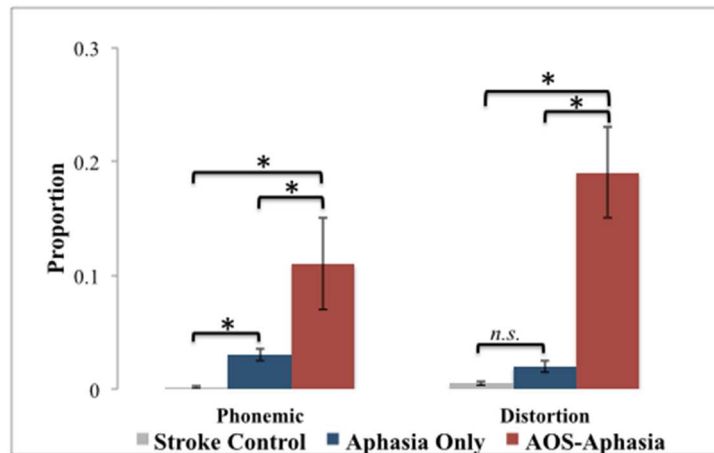


Figure 3.6. Proportions of phonemic errors and distortion errors. Significant differences are indicated by asterisks.

3.2.6 Measures of Speech Fluency. Because a number of different measures have been adopted to measure speech fluency, three measures were analyzed here – mean words per minute (WPM), mean syllabic rate (syll/sec) and mean speech productivity.

The three fluency measures were transformed to standard z-scores and analyzed using a 3x3 (group x fluency measures) MANOVA.⁷ There was a main effect of group, $F(6, 108)=10.15, p<0.001$; partial $\eta^2=0.37$. Follow-up univariate ANOVAs were significant for all three fluency measures, as follows: WPM: $F(2, 55)=27.07, p<0.001$, partial $\eta^2=0.50$; syllabic rate: $F(2, 55)=31.84, p<0.001$, partial $\eta^2=0.54$; and speech productivity: $F(2, 55)=25.21, p<0.001$, partial $\eta^2=0.48$.

Bonferroni corrected post-hoc comparisons showed that the Stroke Control group had higher standardized fluency scores, across all three measures, when compared to both the Aphasia Only and AOS-Aphasia groups (mean difference values=1.84-1.97, $p<0.001$ for all measures). The Aphasia Only group had greater fluency scores for syllabic rate

⁷ Although the z-scored syllabic rate and WPM violated the assumptions of normality for the Stroke Control group, MANOVA analyses are arguably robust to violations of normality (Ito, 1980).

(mean difference=0.87) and mean WPM only (mean difference= 0.81; $p<0.001$ for both). Speech productivity scores did not differ significantly between the Aphasia Only and AOS-Aphasia groups (mean difference=0.52, $p=n.s.$). Figure 3.7 presents standard scores for all fluency measures.

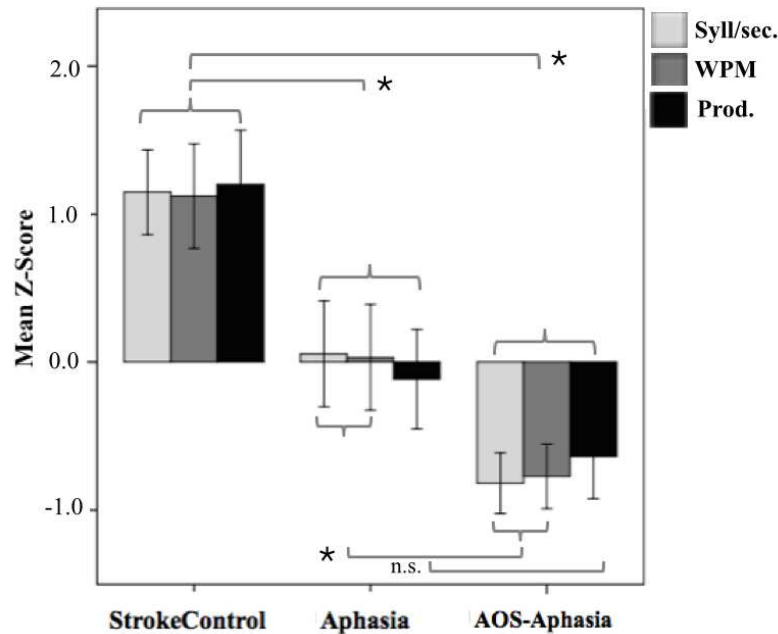


Figure 3.7. Comparison of speech fluency measures (based on standard scores). Significant group differences are indicated by asterisks.

3.2.7 Envelope Modulation Spectrum. Inspection of each participant's EMS plots calculated over the 20-second intervals revealed similar patterns; therefore, EMS variables were averaged across the three 20-second intervals for further analysis. Mean EMS plots for each group are presented in Figure 3.8. Less variability is evident across the 1-32 Hz frequency bands for the AOS-Aphasia group, indicated by flatter peaks across the lower frequencies. The Aphasia Only group trended towards a more defined peak at 4 Hz, a trend more evident for the Stroke Control group. Group differences for

the EMS variables selected for analysis in this study (i.e., peak frequency; peak amplitude; and band prominence for the 1, 2, 4, 8, 16 and 32 Hz bands) are discussed in the sections that follow.

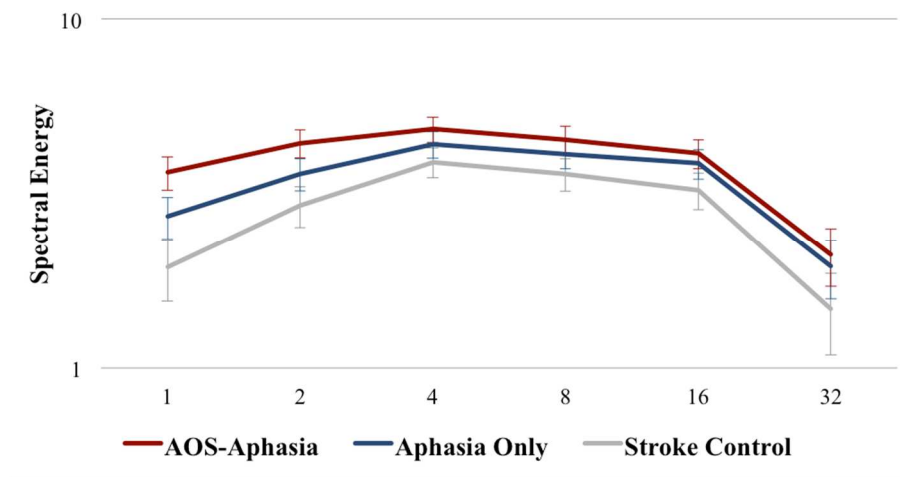


Figure 3.8. Amplitude modulation spectra for each group. The x-axis represents the frequency band measured, in harmonic intervals from 1-32 Hz.

Peak Frequency and Peak Amplitude. Peak frequency, a measure of the frequency that “dominates” production (Liss et al., 2010), was investigated with raw peak frequency values (i.e., selection of the octave band with the highest peak) and the absolute deviation in peak frequency (i.e., the deviation in peak frequency values from the selected 4 Hz reference). Both variables were investigated with non-parametric Kruskal-Wallis ANOVAs. There were no main effects of group for either comparison ($p>0.05$). However, as presented in Figure 3.9, there is a notably greater range in peak frequency and peak frequency deviation values for the AOS-Aphasia group. This range was significantly greater in the AOS-Aphasia group compared to those without AOS (collapsed across the Stroke Control and Aphasia Only groups; peak frequency: Moses span, trimmed for outliers=45, $p<0.01$; deviation in peak frequency: Moses span,

trimmed for outliers=43, $p=0.001$). Direct comparisons between the Aphasia Only and AOS-Aphasia groups were also statistically significant (Moses span, trimmed for outliers, all $p<0.05$). There were no differences between the Stroke Control and Aphasia Only group.

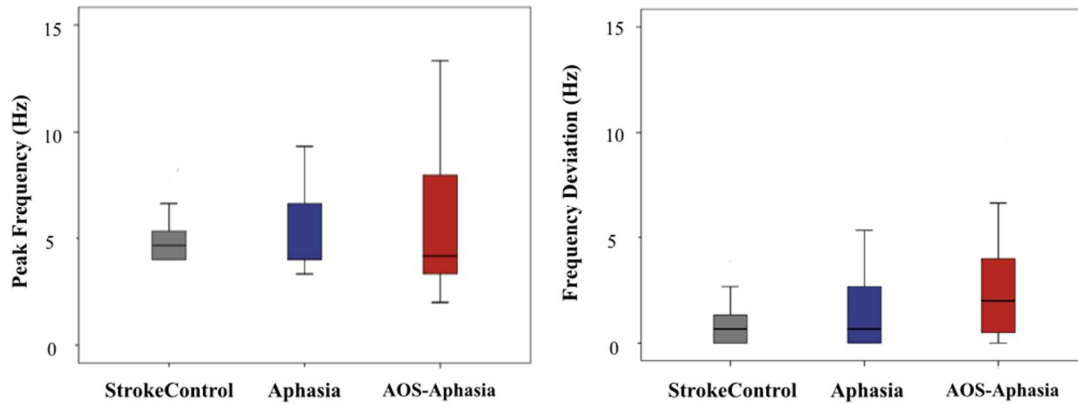


Figure 3.9. Peak frequency and deviation for each group. Panel A presents mean peak frequency, and Panel B presents mean peak frequency deviation from the 4 Hz reference. There were no statistically significant group differences in mean peak frequency for either comparison, but the ranges in peak frequency and deviation were greater in the AOS-Aphasia group compared to those without AOS.

Peak amplitude is a measure of the energy in the peak frequency range; essentially, peak amplitude provides information regarding the amount of energy present at the peak frequency. A Kruskal-Wallis ANOVA revealed a significant group effect in peak amplitude: $\chi^2(2)=10.43$, $p<0.01$. Post-hoc Mann Whitney U-tests show that the AOS-Aphasia and Aphasia Only groups had significantly greater peak amplitude compared to the Stroke Control group (AOS-Aphasia vs. Stroke Control, $U=43$, $z=3$, $p<0.01$; Aphasia Only vs. Stroke Control, $U=72$, $z=-2.53$, $p=0.01$). There were no significant differences between the AOS-Aphasia and Aphasia Only groups ($p>0.05$). Mean peak amplitude is presented in Figure 3.10.

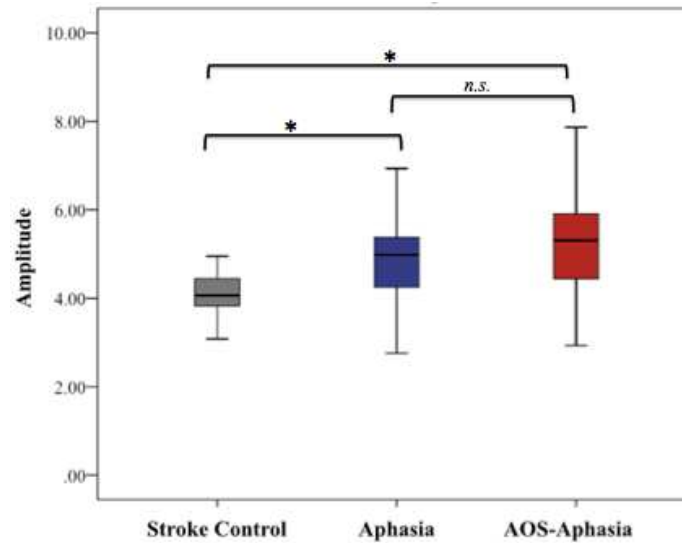


Figure 3.10. Mean peak amplitude for each group. Significant differences in peak amplitude are indicated by asterisks.

Band Prominence. Frequency band-specific EMS variables were investigated using a 3x6 (group x frequency) MANOVA⁸. Frequency bands representative of prosodic (1 Hz, 2 Hz), syllabic (4 Hz, 8 Hz), and phonetic production (16 Hz, 32 Hz) were chosen for comparison. As evident from the overall modulation spectrum presented in Figure 3.8, the AOS-Aphasia group had numerically higher energy across the entire spectrum. Therefore, to compare frequency bands across groups, a measure of band prominence was used, where energy in each individual band was divided by the sum of the energy across all frequencies (Carbonell et al., 2015; Liss et al., 2010).

The main effect of group was significant, $F(12, 94)=3.12$, $p<0.005$, partial $\eta^2=0.29$. Follow-up univariate ANOVAs were significant for all frequency bands except for 32 Hz. Results from the univariate ANOVAs are presented in Table 3.1. Bonferroni-

⁸ A note regarding assumptions. All variables included in this analysis were normally distributed except for the distribution for the 1 Hz band prominence for the Aphasia Only group and 4 Hz band prominence for the Stroke Control group. One univariate outlier from the AOS-Aphasia group was included in this analysis. Linear relationships between variables were rather weak. Despite a potential loss of power, a MANOVA was used to evaluate the EMS band prominence variables.

corrected post-hoc comparisons revealed significant differences in frequency bands between the AOS-Aphasia and Aphasia Only groups (1 Hz, 2 Hz, 8 Hz, 16 Hz) and the AOS-Aphasia and Stroke Control Groups (1 Hz, 4 Hz, 8 Hz, 16 Hz bands). There were no significant differences between the Stroke Control and AOS-Aphasia groups ($p>0.05$ for all pairwise comparisons). Table 3.2 presents a summary of pairwise comparisons.

Table 3.1

Results from univariate ANOVAs for each frequency band

Frequency Band	<i>F</i> (2,52)	η^2	<i>p</i>
1 Hz	14.41	0.36	<0.001
2 Hz	4.30	0.14	<0.05
4 Hz	5.04	0.16	<0.01
8 Hz	6.50	0.20	<0.005
16 Hz	5.48	0.17	<0.01
32 Hz	1.17	0.04	n.s.

Table 3.2

Pairwise comparisons for band prominence values between 1-32 Hz

Comparisons are presented for the AOS-Aphasia group relative to the Stroke Control and Aphasia Only groups. There were no significant differences between the Stroke Control and Aphasia Only groups for any of the frequency bands evaluated.

Frequency Band	AOS-Aphasia v. Stroke Control	AOS-Aphasia v. Aphasia Only
1 Hz	0.039, $p<0.001$	0.026, $p=0.001$
2 Hz	0.017, $p=n.s.$	0.016, $p<0.05$
4 Hz	-0.022, $p<0.01$	-0.011, $p=n.s.$
8 Hz	-0.021, $p<0.005$	-0.014, $p<0.05$
16 Hz	-0.019, $p<0.05$	-0.017, $p<0.05$
32 Hz	0.00, $p=n.s.$	-0.01, $p=n.s.$

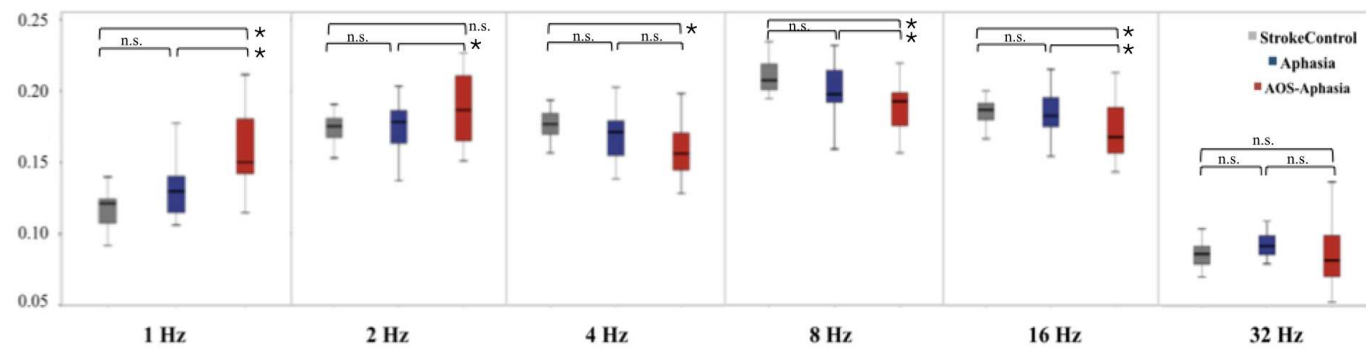


Figure 3.11. Band prominence for each frequency band within the 1-32 Hz range. Significant differences are indicated by asterisks.

A final analysis was conducted to determine if specific speech production behaviors were driving these results. A series of correlation analyses showed that the deviation in peak frequency and 2 Hz band prominence were significantly correlated with only the three rate measures (WPM, syllabic rate, productivity; $-0.32 < r_s > -0.43$, all $p < 0.005$), and 32 Hz was not correlated with any other variables obtained here. All remaining correlations are presented in Table 3.3.

Table 3.3

Correlation coefficients for the EMS measures and speech production deficits

Correlations that survive Bonferroni correction ($p \leq 0.01$) are marked with one asterisk (*). Correlations that survive Bonferroni correction when controlled for aphasia severity (WAB AQ scores) are indicated by two asterisks (**).

Behavior	Peak Amp.	1 Hz Band	4 Hz Band	8 Hz Band	16 Hz Band
AOS Severity	-0.27	0.58*	-0.47*	-0.43*	-0.34
Aphasia Severity	0.43*	-0.45*	0.45	0.27	0.05
nPVI-V	0.11	-0.35*	0.25	0.50**	0.20
VOT-SD_{voiced}	0.05	0.31	0.03	-0.26	-0.20
VOT-SD_{voiceless}	-0.23	0.38*	-0.30	-0.29	-0.01
WPM	0.43	-0.66*	0.54**	0.57**	0.26
Productivity	0.48*	-0.52*	0.60**	0.43*	0.15
Syllabic Rate	0.48*	-0.68*	0.59**	0.57**	0.28
Dist. Errors	-0.31	0.53*	-0.34	-0.42**	-0.30**
Phon Errors	-0.45*	0.49*	-0.52**	-0.41**	-0.14

Across the EMS variables selected here, the direction of significant correlations differed as a function of increasing frequency. This pattern was seen for the relationship between the three fluency measures and 1 Hz, and 4 Hz band prominence, where greater deviations in 1 Hz band prominence were related to decreased fluency (see negative correlations presented in Table 3.1), and greater band prominence at 4 Hz was associated with more fluent speech. Notably, none of the correlation coefficients surpassed $r = +/-$

0.68, suggesting that these measures may not be completely redundant with speech fluency (see also Liss et al., 2010).

Interestingly, correlations with distortion errors, which arguably distinguish apraxic errors from phonemic errors (Cunningham et al., 2015), showed a U-shaped pattern. As presented in Figure 3.12, higher energy at the 1 Hz band was associated with more sound errors (both distortion and phonemic errors). Conversely, mean sound errors and 4 Hz band prominence were inversely correlated, where more energy at 4 Hz was related to fewer sound errors. The strength of this relationship decreased for the higher frequency bands (16-32 Hz), although only the correlation between 16 Hz and distortion errors was significant (uncorrected, $p < 0.05$). This same pattern was seen for ASRS AOS severity scores.

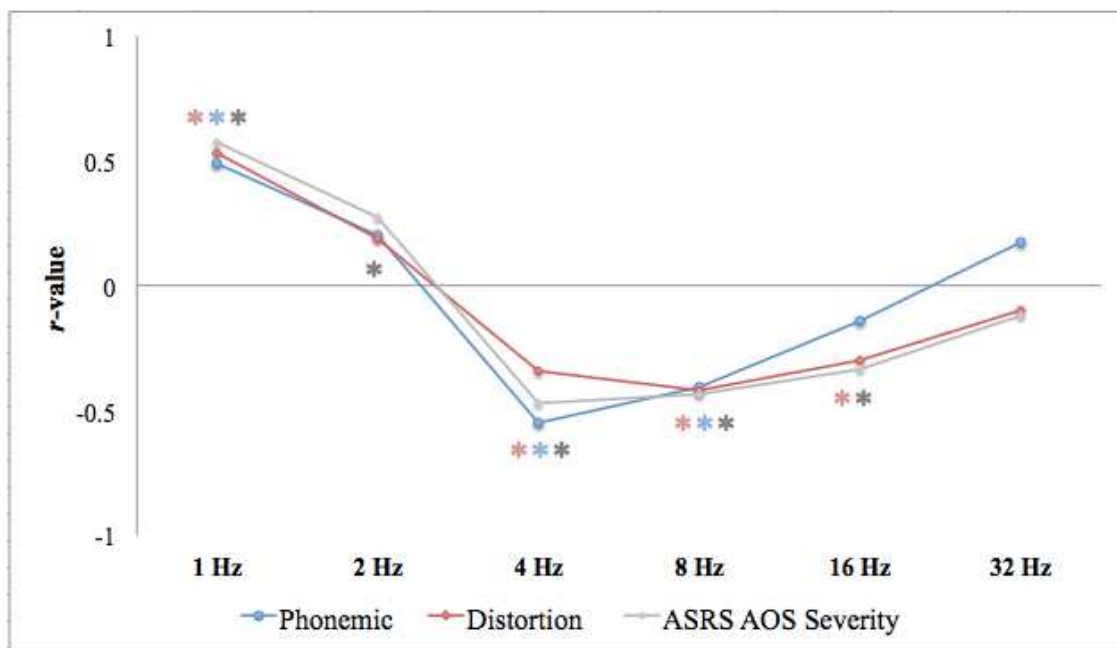


Figure 3.12. Correlations between speech sound errors, ASRS AOS severity and band prominence. Asterisks indicate significant correlations between each frequency band and phonemic errors (blue), distortion errors (red) and ASRS AOS severity (grey).

3.2 SUMMARY OF BEHAVIORAL MEASURES

Overall, comparisons for each measure show significant group differences in stress pattern in connected speech (relative durations of successive vowel segments, nPVI-V), timing of articulatory coordination (VOT variability), the proportion of phonemic and distortion errors to total words produced, speech fluency (WPM, speech productivity, and syllables/second), and several of the measures obtained from the EMS analysis. The individuals in the AOS-Aphasia group performed significantly different than those with no speech/language impairment on all measures (except 2 and 32 Hz band prominence; Figure 3.11), and compared to the Aphasia Only group, those with AOS demonstrated significant differences in all measures obtained from the connected speech samples, except for speech productivity (Figure 3.7) and band prominence in the 4 and 32 Hz frequency bands (Figure 3.11). A complete summary of all means, standard deviations, and correlations can be found in Appendices F and G.

It should be noted that the AOS-Aphasia group had significantly lower WAB AQ scores (i.e., more severe aphasia). Accordingly, correlation analyses were conducted to determine if relationships exist each of the measures obtained here and aphasia severity, which could explain some of the current findings. Aphasia severity was correlated with three of the EMS variables (1 Hz and 4 Hz band prominence, peak amplitude), the three fluency measures (syllabic rate, WPM, speech productivity), proportion of phonemic errors, and proportion of distortion errors. WAB AQ scores were not correlated with the remaining measures (i.e., nPVI-V, VOT-SD_{voiced}, VOT-SD_{voiceless}, peak frequency, 2 Hz, 8 Hz 16 Hz or 32 Hz band prominence). Correlations coefficients are presented in Table

3.4 below. The relationships between aphasia severity and these variables will be addressed further in the LDA and neuroimaging analyses.

Table 3.4

Correlation coefficients between WAB AQ scores and all study measures.

Correlation coefficients flagged by an asterisk indicate a significant correlation at the Bonferroni-corrected level of significance, $p < 0.003$ (i.e., 0.05/15 comparisons).

Correlations with Aphasia Severity	
Behavior	<i>r</i>-value
nPVI-V	0.22
VOT-SD_{voiced}	-0.09
VOT-SD_{voiceless}	-0.27
Phonemic Errors	-0.76*
Distortion Errors	-0.48*
Peak Amplitude	0.43*
Peak Frequency	-0.27
1 Hz Band	-0.45*
2 Hz Band	-0.10
4 Hz Band	0.45*
8 Hz Band	0.27
16 Hz Band	-0.12
32 Hz Band	-0.29
Syllabic Rate	0.59*
Productivity	0.66*
WPM	0.55*

3.3 CLINICAL CLASSIFICATION USING LINEAR DISCRIMINANT ANALYSIS

From the above analyses, it is apparent that individuals with AOS demonstrate performance that is significantly different when compared to individuals with brain damage but without any communication impairment, and in most cases, when compared to individuals with aphasia only. However, a large number of variables was evaluated here, and the variable with the greatest discriminative weight in AOS classification remains uncertain from the above analyses. Therefore, the linear discriminant function was used to determine how accurate these measures are in predicting AOS (defined by

ASRS criteria; Strand et al., 2014), and which measures have the greatest discriminative weight. Although normality is generally assumed in LDA, it has been argued that this assumption can be relaxed with sample sizes greater than 50 (Pohar, Blas, & Turk, 2004).

AOS classification was tested on a series of binary comparisons as follows: 1) the AOS-Aphasia group vs. all other participants, 2) the AOS-Aphasia group vs. the Aphasia Only group, and 3) the AOS “subgroups” based on nPVI-V scores. Because LDA is sensitive to highly correlated predictor variables, care was taken in variable selection (see also Section 2.5.1). Variables were chosen based on 1) theoretical rationale, 2) magnitude of correlations with other variables, and 3) magnitude of correlations with aphasia severity (WAB AQ scores).

According to these criteria, nPVI-V and VOT variability⁹ from the connected speech samples were selected based on prior literature (Auzou et al., 2000; Ballard et al., 2016; Seddoh et al., 1996; Vergis et al., 2014). Next, 4 Hz and 16 Hz energy were selected because neither of these variables demonstrate significant correlation with nPVI-V or VOT-SD_{voiced}, and neither demonstrate correlations with aphasia severity beyond r values of 0.50. Importantly, these variables are reportedly reflective of syllabic (4 Hz) and phonetic (16 Hz) levels of production (Crouzet & Ainsworth, 2001; Ghitza, 2011, 2013; Ghitza et al., 2012; Ghitza & Greenberg, 2009; Poeppel, 2003), which have been debated as the source of apraxic deficits. Finally, although the proportions of phonemic errors and distortion errors were highly correlated within themselves ($r_s=0.62$) and with aphasia severity (phonemic errors: $r_s=-0.76$; distortion errors: $r_s=-0.48$), these variables

⁹ VOT variability for voiced stop consonants (VOT-SD_{voiced}) was chosen here due to prior work suggesting that VOT for voiceless stop consonants is more susceptible to other production factors, including speech rate (Kessinger & Blumstein, 1998).

have been commonly discussed in the aphasia and AOS literature. Therefore, to decrease the number of inter-correlated variables in the model, these error counts were averaged to create a “mean sound errors” variable. Because rating these behaviors can be somewhat subjective, LDA models were run with and without this "mean sound errors" variable to inspect classification accuracy when only objective measures were included. Notably, fluency measures were not included, as it has been suggested that fluency is not a unique feature of AOS (Galluzzi et al., 2015). The results from each model are presented in the sections that follow.

3.3.1 AOS-Aphasia group vs. All participants without AOS. The model that included all of the aforementioned variables (i.e., nPVI-V, VOT-SD_{voiced}, 4 Hz and 16 Hz band prominence, mean sound errors) was 90.7% accurate in classification (percentage reflects leave-one out cross-validation, LOOCV; $p=6.14 \times 10^{-7}$; sensitivity=85, specificity=94). Visualization of the groups using MDS (Figure 3.13) shows a linear division between those with AOS and those without, with some participants straddling this boundary. The variables with the greatest discriminative weight for AOS classification included 16 Hz energy, and VOT-SD_{voiced}. Figure 3.14 (left panel) presents classification weights for all variables; variables with negative weights are predictive of AOS.

The second model analyzed, which excluded the “mean sound errors” variable, was slightly less accurate (85% cross-validated accuracy; $p=7.99 \times 10^{-05}$; sensitivity=80, specificity=88). Classification weights for the remaining variables are presented in Figure 3.14b. Once again, VOT-SD_{voiced} and 16 Hz band prominence demonstrated the greatest discriminative weight in the analysis.

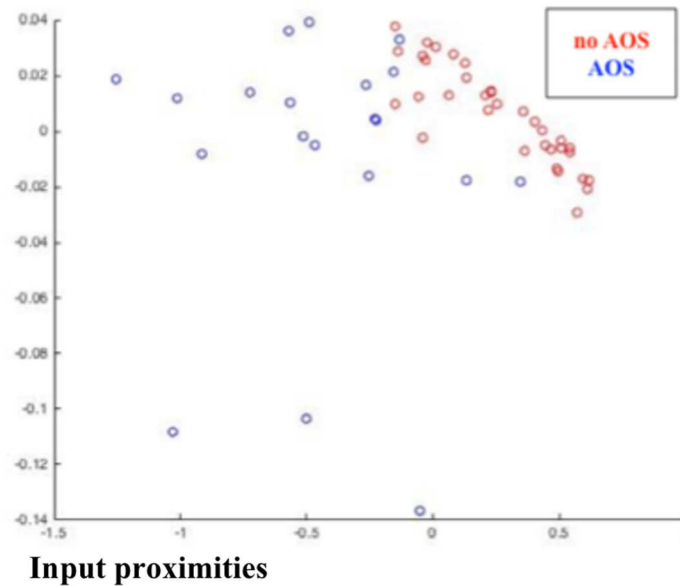


Figure 3.13. MDS plot for the AOS-Aphasia group versus those without AOS (i.e., Stroke Control and Aphasia Only groups combined). Axes reflect input proximities based on variables included in the MDS.

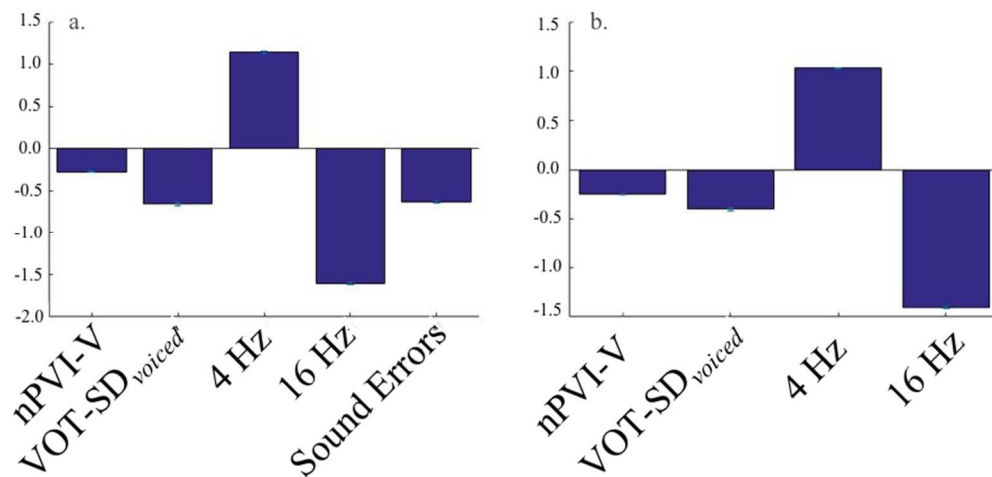


Figure 3.14. Discriminative weights obtained from the two models used to classify participants with AOS from those without AOS. Panel A presents discriminative weights from the model with the highest classification accuracy. Panel B presents discriminative weights from the model that only included objective measures. Performance on measures with a negative weight is indicative of the AOS classification, whereas measures with a positive weight are indicative of no AOS classification.

3.3.2 AOS-Aphasia group vs. Aphasia Only group. Differential diagnosis of AOS is most clinically challenging in the presence of aphasia. Therefore, a model that excluded all individuals without speech/language impairment was tested to determine the accuracy of the aforementioned variables when discriminating between speakers with AOS and concomitant aphasia, from those with aphasia only. When excluding the Stroke Control group, overall classification accuracy decreased to 88.64% ($p < 6.14 \times 10^{-7}$; sensitivity=80, specificity=96), and once again, decreased when sound errors were removed, to 84.09% accuracy ($p < 7.51 \times 10^{-5}$; sensitivity=75, specificity=92). Figure 3.15 presents discriminative weights for both models. Discriminative weights for the first model are presented in panel A, and discriminative weights for the second model are presented in panel B. As displayed in these figures, 16 Hz band prominence, proportion of sound errors, and VOT-SD_{voiced} had the greatest discriminative weight for AOS classification.

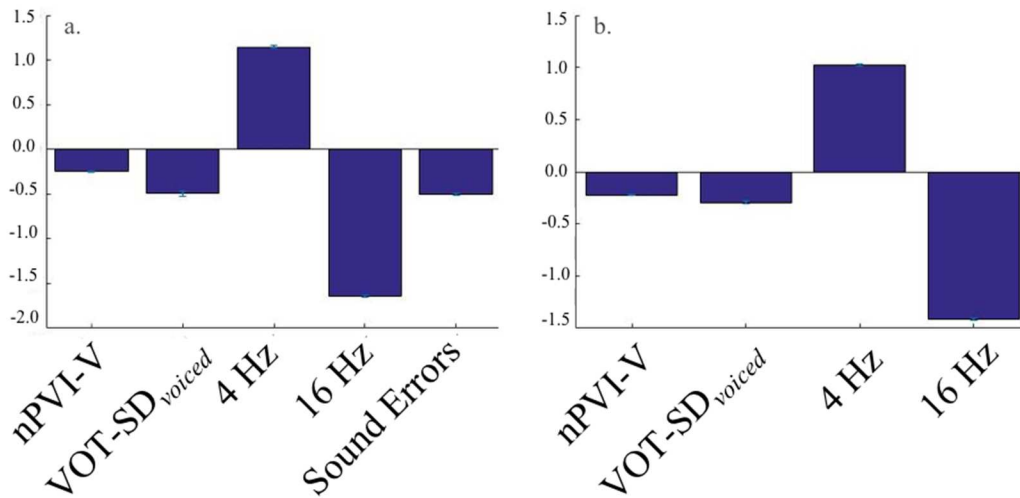


Figure 3.15. Discriminative weights obtained from the models used to classify the AOS-Aphasia participants from the Aphasia Only participants.

3.3.3 AOS “Subtypes.” As seen in Section 3.2.1, some individuals in the AOS-Aphasia group had nPVI-V coefficients that were within one standard deviation of the mean of the Aphasia Group, while others had coefficients beyond one and two standard deviations. To investigate the extent that the AOS-Aphasia can indeed be divided into subgroups, all participants with AOS were assigned to a subgroup based on nPVI-V scores, where individuals were assigned to a “1.5+ nPVI” group if their nPVI-V coefficient was below 1.5 standard deviations from that of the Aphasia Only group. All other participants in the AOS-Aphasia group were assigned to the “within 1.5 nPVI” group. A standard deviation of 1.5 was chosen so groups would be balanced (both groups $n=10$). Classification between these two groups was then based on VOT-SD_{voiced}, 4Hz and 16 Hz band prominence, and the proportion of sound errors. The rationale here is that according to prior work (e.g., Josephs et al., 2013), those with greater rhythmic impairments should present with fewer distortion errors, and vice versa. Additionally, these variables were not significantly correlated with nPVI-V, the measure that was used for group assignment.

Classification accuracy was slightly worse than chance-level (45%, $p=0.59$), suggesting that although there may be subgroups of individuals with AOS based on their nPVI-V scores, they could not be reliably classified based on other production measures obtained here. To determine if one “subgroup” indeed had more sound-level errors than the other (Duffy & Josephs, 2012), distortion errors between the low and high nPVI-V groups were compared using a Mann Whitney U-test. There were no significant group differences ($p>0.05$).

3.3.4 Relationships between objective and subjective measures. Taken together, the LDA results consistently revealed that 16 Hz band prominence, proportion of sound-level errors, and VOT variability for voiced stop consonants have the greatest discriminative weight in the classification of AOS. These top three AOS predictors, and the top predictor for non-AOS classification (i.e., 4 Hz) were entered into a correlation analysis with ASRS items to determine the relationship between perceptual features (i.e., each of the ASRS items) and these objective measures of speech production.

In the first correlation analysis, all items from the first two ASRS sections (Items 1.1-1.6, Items 2.1-2.6) were entered into a correlation with 16 Hz band prominence, sound errors, and VOT-SD_{voiced}. In the second correlation analysis, all remaining items (3.1-3.2, 4.1-4.2) were entered with the same predictor variables, but controlling for aphasia severity (WAB AQ), as these ASRS items can be associated with production errors that occur in aphasia. Correlation results are presented in Table 3.5.

Table 3.5

Correlations between ASRS items and variables with the greatest weight in diagnostic classification

Correlations reported for Items 3.1-4.2 were controlled for aphasia severity (WAB AQ), as behaviors scored in these items can also occur in aphasia. Items flagged by an asterisk survived Bonferroni correction (significant $p < 0.003$). Correlations with the proportion of sound errors is not presented here, as this variable was highly correlated with all ASRS items except for 4.1 and 4.2 ($0.65 < r_s > 0.82$ for Items 1.1-3.2).

ASRS Items	4 Hz Band	16 Hz Band	VOT- SD _{voiced}
Item 1.1 Distorted substitutions	-0.43*	<i>n.s.</i>	0.41*
Item 1.3 Increased distortions with increased length	-0.41*	<i>n.s.</i>	<i>n.s.</i>
Item 1.5 Inaccurate AMRs	-0.46*	<i>n.s.</i>	<i>n.s.</i>
Item 2.3 Sound distortions	<i>n.s.</i>	-0.39*	0.43*
Item 2.6 Lengthened intersegment durations	<i>n.s.</i>	-0.39*	<i>n.s.</i>
Item 3.2 Audible/visible articulatory groping	<i>n.s.</i>	-0.50*	<i>n.s.</i>

As evidenced by these results, 16 Hz energy and VOT-SD_{voiced} correlated significantly with ASRS items pertaining to phonetic-level production, whereas the proportion of sound errors was correlated with all items except for 4.1 (sound/syllable repetitions) and 4.2 (sound prolongations, beyond lengthened segments). Interestingly, 16Hz energy was significantly correlated with Item 2.6 (lengthened intersegment durations), an item reflective of perceptual evaluation of speech rhythm/rate. In the current sample, Item 2.6 was highly correlated with sound distortions (Item 2.3; $r_s=0.83$), suggesting that those with sound distortions will generally demonstrate altered segmental production, hence the correlation with these items and 16 Hz energy. Taken together, the selectivity of these correlations suggests that 16Hz energy and VOT-SD_{voiced} may serve as more sensitive measures of phonetic production than measures of sound errors; however, further study is needed to verify this claim.

3.4 NEUROIMAGING RESULTS: VLSM AND CONNECTOME ANALYSES

VLSM analyses were used to identify patterns of damage predictive of each behavior. Because aphasia severity was correlated with some of the variables included here, multivariate VLSM analyses were conducted with aphasia severity entered as a co-factor in all analyses¹⁰. Additional multivariate VLSM analyses were conducted with the EMS variables, as described below. For variables that did not survive permutation thresholding when aphasia severity was included as a co-factor, or that did survive but with a small (<30 voxels) statistical map, a follow-up univariate analysis was conducted

¹⁰ Notably, lesion size did not differ between the Aphasia Only and AOS-Aphasia groups, $t(41)=1.53$, $p=.13$, although both groups had significantly greater lesion volumes when compared to the Stroke Control group [A.O. vs. S.C.: $t(34)=2.84$, $p<.01$; A.-A. vs. S.C.: $t(31)=5.67$, $p<.001$).

that included only the behavior in question. Significant statistical maps are presented in Figure 3.16, and Table 3.6 lists the analyses that survived permutation thresholding.

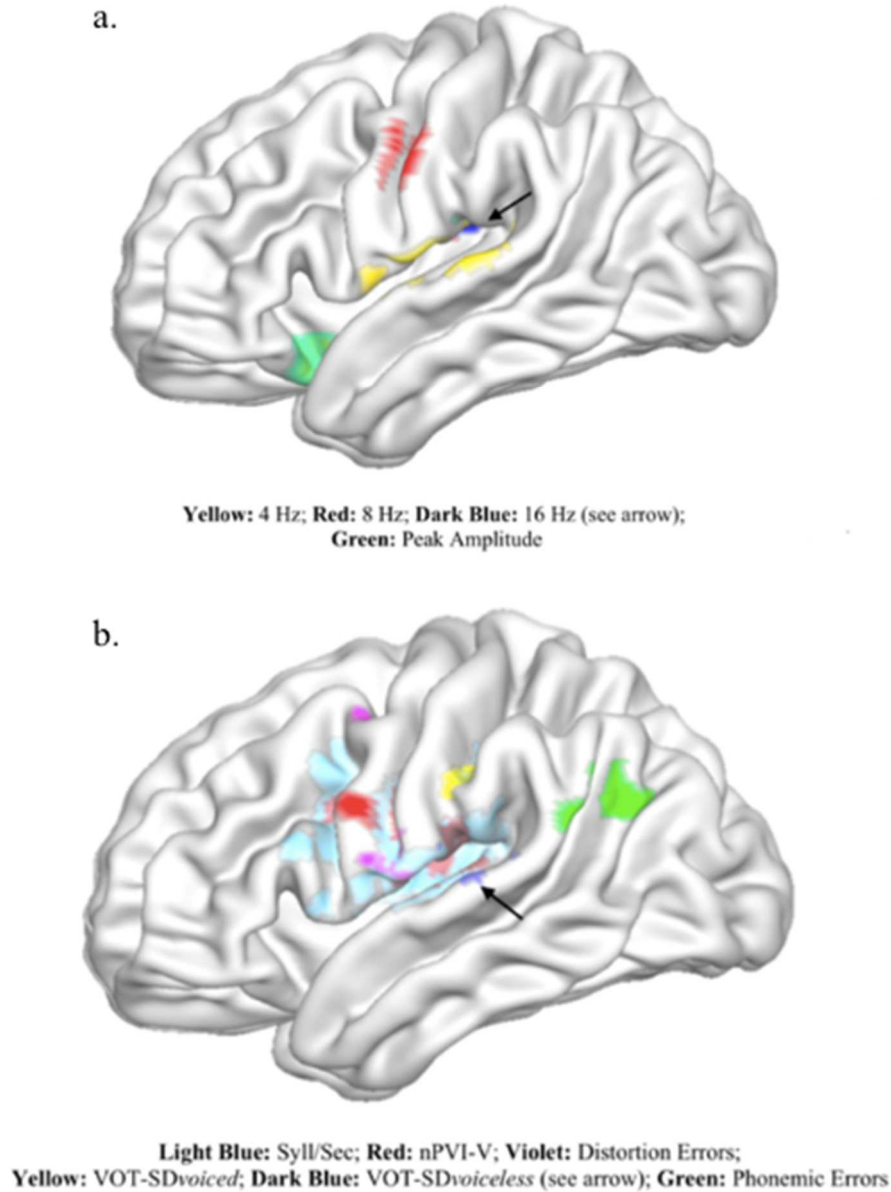


Figure 3.16. Results from significant VLSM analyses. Panel A presents significant statistical maps from the EMS variables. Panel B presents significant VLSM results from the remaining behavioral variables. All variables presented here were analyzed with aphasia severity as a nuisance regressor, except for 4 Hz band prominence (Panel A, yellow) and peak amplitude (panel A, green) as these variables did not survive permutation thresholding when aphasia severity was included.

Table 3.6

Regions implicated in VLSM analyses.

This table presents the brain regions that comprise the statistical maps that emerged from the significant VLSM analyses. Unless otherwise indicated, these results are from the multivariate (MV) analysis that controlled for aphasia severity. The size (in voxels) of each statistical map is presented in parentheses alongside each variable. The first column presents the regions involved in each statistical map. The percentage of each region implicated in the statistical map is presented in the second column, and the third column states the percentage of the region implicated in each statistical map. Regions listed here comprise at least 5% of the total statistical map.

VLSM Results		
Region	% Statistical map coverage	% Region Implicated
Peak Amplitude (316 voxels, univariate only)		
Ins	49.05	2.44
LFOG	30.06	1.06
pSTG	5.38	0.17
4 Hz Band (10 voxels, MV)		
sCR	100	0.12
4 Hz Band (366 voxels, univariate)		
PrCG	28.14	0.32
PoCG	10.03	0.22
STG	15.85	0.37
Ins	10.93	0.63
pSTG	6.56	0.26
IFGpo	5.19	0.23
8 Hz Band (167 voxels, MV)		
PrCG	43.14	0.22
PoCG	29.94	0.16
sCR	16.17	0.36
16 Hz Band (21 Voxels, MV)		
SMG	66.67	0.7
PrCG	33.33	0.21
nPVI (1316 voxels, MV)		
PoCG	44.23	1.9
SLF	18.69	3.14
PrCG	11.7	0.47
MFG	10.26	0.52
SMG	7.98	

VLSM Results, Continued		
Region	% statistical map coverage	% Region Implicated
Phonemic Errors (146 voxels, MV)		
AG	51.37	0.36
MOG	36.99	0.17
pTR	6.85	0.17
Distortion Errors (61 voxels, MV)		
MFG	40.98	0.1
PrCG	32.79	0.06
PoCG	8.2	0.02
Syllabic Rate (14755 voxels, MV)		
PrCG	24.1	10.9
PoCG	14.77	7.13
SMG	13.98	10.09
SLF	11.97	22.56
sCR	11.94	21.9
MFG	7.08	4.01
IFGpo	6.85	12.22
VOT-SD_{voiceless} (31 voxels, MV)		
SMG	10	0.02
PoCG	6.67	0.01
VOT-SD_{voiced} (36 voxels, MV)		
SMG	97.22	2.78

Abbreviations: AG: angular gyrus; Ins: Insula; IFGpo: inferior frontal gyrus *pars opercularis*; LFOG: left fronto-orbital gyrus; pSTG: posterior superior temporal gyrus; MFG: middle frontal gyrus; MOG: middle occipital gyrus; PrCG: precentral gyrus; PoCG: post-central gyrus; pTR: posterior thalamic radiation; sCR: superior corona radiata; SLF: superior longitudinal fasciculus; SMG: supramarginal gyrus; STG: superior temporal gyrus

The statistical maps for each of the above variables were inspected to determine the extent that each are represented by unique patterns of cortical damage. Despite a large degree of overlap in nPVI-V and syllabic rate (1020 shared voxels), 296 voxels were unique to nPVI-V, with over half of this unique cluster located in the PrCG (61.59% of the statistical map).

Except for the EMS variables, none of the remaining statistical maps shared more than 10 voxels with any other variable. Accordingly, the variables obtained from the EMS analysis were inspected for overlap between each frequency band and the other behavioral variables. Aside from overlap between 4 Hz and 16 Hz band prominence, and 4 Hz and syllables/second, none of the EMS variables shared considerable overlap with any of the other variables when entered aphasia severity was included as a nuisance regressor.

To inspect the relationship between 4 Hz and 16 Hz, a multivariate analysis with these two variables revealed two significant clusters – a larger cluster (2,564 voxels) predictive of 4Hz energy, and a smaller (122 voxels) cluster predictive of 16Hz energy. As displayed in Figure 3.17, the cluster predictive of 16Hz energy overlapped highly with that of 4Hz energy (91 of 122 voxels shared with 4Hz energy, region indicated by arrow). This overlapping cluster was localized to the SMG (61.54%), SCR (27.47%), PrCG (7.69%) and SLF (3.30%).

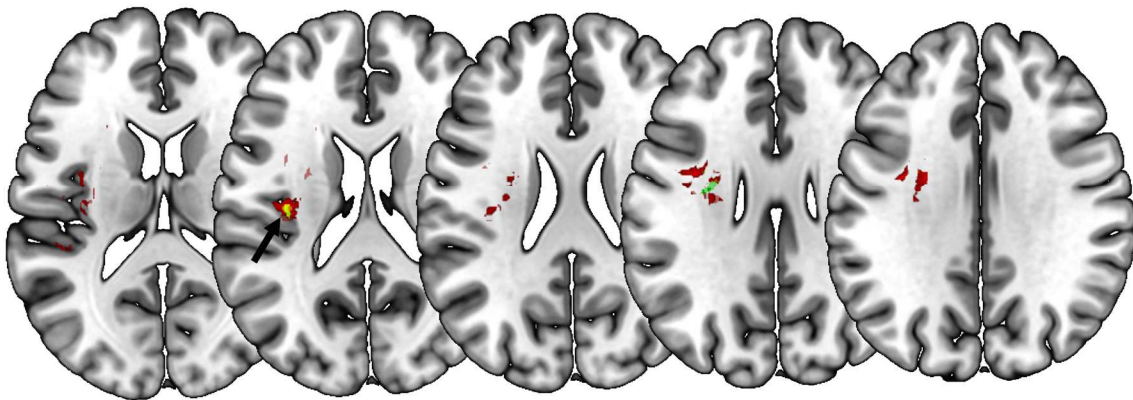


Figure 3.17. Significant regions predictive of 4 Hz (red) and 16 Hz (green) band prominence. The region of overlap between 16Hz and 4Hz is presented in yellow (and indicated by the arrow).

Next, the cluster predictive of 4Hz band prominence (obtained from the univariate analysis) was compared to the cluster predictive of syllabic rate (essentially, the rate of vowels produced per second). 266 voxels were shared between these two measures, with over half of this overlap located in the PrCG and PoCG (70.8% collectively), and some shared involvement of the IFGpo (8.41%) and the STG (7.5%). Inspection of the statistical map unique to the 4 Hz cluster was located in the STG (29.29%), insula (25.71%), pSTG (17.14%), and unique portions of the PrCG (5.7%). Additional regions (external capsule, IFGpt, posterior insula, PoCG, and the pole of the STG) comprised less than 10% of this map. The results from the VLSM analysis with 4 Hz band prominence and speech fluency (measured here as syllabic rate) were analogous to behavioral analyses that included these measures. That is, these results show that lesion damage predictive of syllables/second and 4 Hz band prominence overlaps, but that neuroanatomically, these measures are not completely redundant.

3.4.1 Connectome Analysis. Connectome data were analyzed as follows: 1) structural connectomes were compared between each group, 2) a whole brain DTI analysis was conducted to identify the connections that were predictive of each variable, and 3) an exploratory analysis correlated each behavior with fiber counts between ROIs implicated in models of speech production.

Comparisons of structural connectomes. To compare connectomes across each group, a multidimensional array with dimensions 189x189xZ was created for each group, where x =each JHU region, y =fiber counts between each region, and z =the number of participants in each group. A series of independent groups t-tests were run, where the distribution of every cell in one multidimensional array was compared to that of its

corresponding cell in the other group's matrix. First, comparisons were made across all 189 JHU regions, resulting 189x189 comparisons (i.e., 35,721 t-tests). Because fiber count matrices are constructed to report fiber counts between regions A-B, and the reverse, regions B-A, there are essentially 17,860.5 non-redundant comparisons¹¹. Therefore, for the purpose of multiple comparisons, the Bonferroni-corrected level of significance was set to $p=0.05$ divided by 17,860.5, resulting in a p value of 2×10^{-6} .

Across all t-tests, the only connections to survive this strict level of correction came from fiber count comparisons between the AOS-Aphasia and Stroke Control groups. The AOS-Aphasia group had significantly fewer fiber count connections between the L IFGpo and L PrCG ($p=2.66 \times 10^{-7}$) and the L superior corona radiata (L sCR) and L SMG ($p=2.45 \times 10^{-6}$).

Because not all 189 regions included in the previous analysis are implicated in speech production, a second series of t-tests compared fiber count connections between 62 regions that have been implicated in contemporary speech production models (e.g., DIVA, HSFC). Connectome maps for these 62 regions are presented in Figure 3.18, and a complete list of regions can be found in Appendix E. Once again, the comparison between the AOS-Aphasia and Stroke control groups was the only to yield connections that survived Bonferroni correction when $p=2 \times 10^{-6}$ – the two reported in the previous paragraph (L IFGpo-L PrCG, L sCR-L SMG), as well as fiber count connections between the LIFGpt and LPrCG ($p=5.27 \times 10^{-6}$). Because this analysis was exploratory, subsequent between-groups comparisons were based on uncorrected significance level of $p=0.01$.

¹¹ There are an uneven number of regions in the JHU atlas due to inclusion of the III and IV ventricles as a single ROI; hence, the remainder of 0.5 when counting non-redundant comparisons

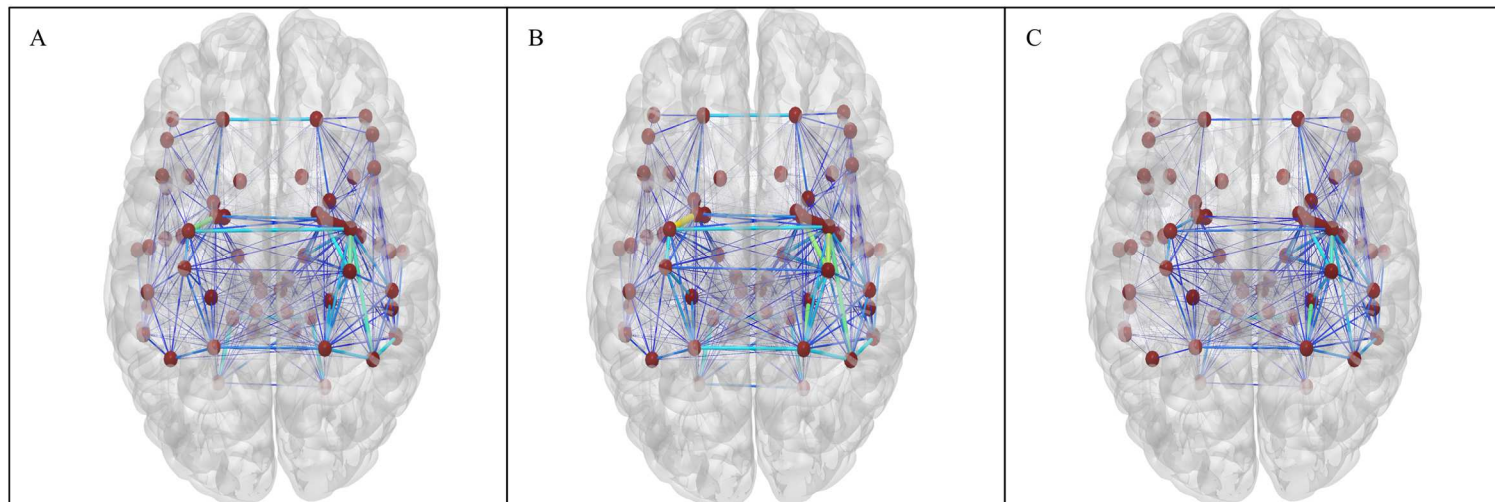


Figure 3.18. Mean fiber count plots for the interregional white matter connections selected for further inspection. Panel A: Stroke Control group; Panel B: Aphasia Only group; Panel C: AOS-Aphasia group.

AOS-Aphasia vs. Stroke Control. At an uncorrected level of significance (i.e., $p < 0.01$), fiber counts between an additional 86 pairs of regions were significantly different between the AOS-Aphasia and Stroke Control groups. Generally, these include connections that involved the L and R IFGpo, L and R IFGpt, L and R insula, L MTG, L STG, L and R corona radiata, L caudate and putamen, and L and R cerebellum. Across all connections the AOS-Aphasia group had fewer fiber count connections between regions. Due to the large number of significantly different connections between these groups, subsequent comparisons were restricted to the fourteen connections that were significantly different between the AOS-Aphasia and Aphasia Only groups (described immediately below).

AOS-Aphasia and Aphasia Only Comparison. There was a significant (uncorrected) difference in fiber counts between fourteen pairs of regions. These regions are presented in Table 3.7, along with mean fiber counts for each group. The fourteen connections are also displayed in Figure 3.19. In all cases, the AOS-Aphasia group had fewer fiber counts between in these connections than either the Stroke Control or Aphasia Only groups.

The following sections present results from analyses correlating fiber count connections to each of the speech production measures.

Table 3.7

Fiber counts derived from structural connectomes

This table presents regions where the AOS-Aphasia group demonstrated significantly fewer fiber count connections when compared to the Aphasia Only group. Each column presents fiber counts for each group.

Fiber Counts by Group			
Regions	Stroke Control	Aphasia Only	AOS-Aphasia
LPoCG-LaCR	3.55×10^6	1.83×10^6	5.59×10^5
LPrCG-LaCR	6.28×10^6	3.70×10^6	9.85×10^5
LPoCG-LAG	4.59×10^7	2.24×10^7	1.49×10^6
RPrCG-LCaud	1.75×10^5	2.57×10^5	3.50×10^3
LSTG-LGP	6.68×10^5	1.03×10^6	4.60×10^5
LPrCG-LIFGpo	4.00×10^7	2.58×10^7	7.65×10^6
LCaud-LpCR	3.43×10^5	4.34×10^5	3.04×10^4
LIFGorb-LPoCG	1.74×10^6	3.48×10^5	6.37×10^4
LAG-LsCR	2.20×10^7	4.59×10^6	4.54×10^5
LCaud-LsCR	1.50×10^6	9.76×10^5	7.74×10^4
LSMG-LsCR	1.33×10^7	6.16×10^6	2.35×10^6
LPrCG-LSMG	4.59×10^7	3.26×10^7	4.16×10^6
LCaud-RsCR	1.61×10^6	1.56×10^5	2.59×10^4
LAG-RIFGpo	2.17×10^5	7.59×10^4	2.23×10^4

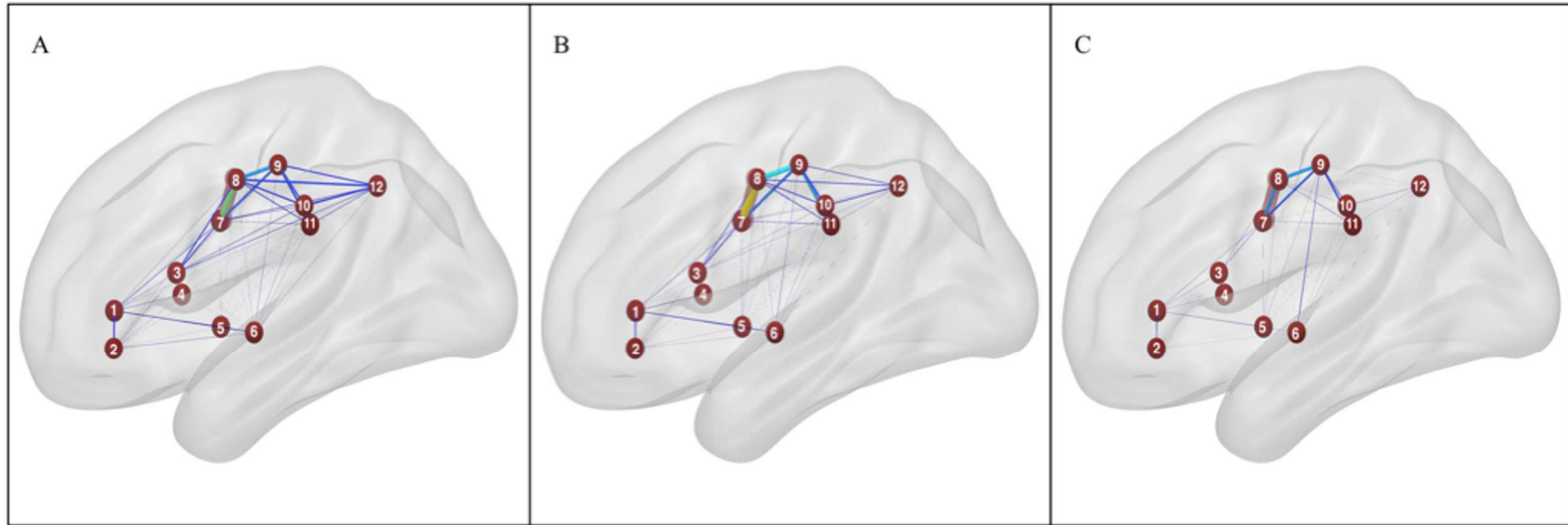


Figure 3.19. Interregional fiber count connections used for further connectome analyses. These connections were chosen because they were found to be significantly different between the Aphasia Only and AOS-Aphasia groups (when $p < 0.01$). Panels are as follows: A: Stroke Control Group, B: Aphasia Only Group, C: AOS-Aphasia group (C). Warmer colors indicate higher fiber counts between region pairs. Regions are as follows: 1) aCR; 2) IFGorb; 3) Caudate; 4) IFGpo; 5) GP; 6) STG; 7) sCR; 8) PrCG; 9) PoCG; 10) SMG; 11) pCR; 12) AG. Connections to right hemisphere regions not pictured here.

3.4.2 Whole Brain DTI Analyses. When controlling for aphasia severity, the only DTI analyses to survive thresholding were the analyses of 1 Hz and 8 Hz band prominence. Follow-up univariate analyses for the remaining behaviors were significant only for the proportion of distortion errors and proportion of phonemic errors. All results from the DTI analyses are presented in Table 3.8.

Table 3.8

Results from significant whole brain DTI analyses

Behaviors marked with an asterisk (*) were those what survived permutation thresholding when aphasia severity was included as a co-factor.

Behavior	Connections	Z-Score
1 Hz Band Prominence*	R MOG-L Lingual	$z=7.16$
	R STGpole-R ACC	$z=7.10$
	R MOG-L Cerebellum	$z=7.16$
8 Hz Band Prominence*	R SFG-R PCG	$Z=3.93$
Phonemic Errors	L IFGpo-L PrCG	$z=-3.11$
	L Thalamus-L CP	$z=-3.22$
Distortion Errors	L MFG-L PrCG	$z=-2.96$
	L IFGpo-L PrCG	$z=-2.84$
	L PrCG-L SMG	$z=-2.72$

Abbreviations are as follows: MOG: middle occipital gyrus; STG: superior temporal gyrus; ACC: anterior cingulate cortex; SFG: superior frontal gyrus; PCG: posterior cingulate gyrus; IFGpo: inferior frontal gyrus pars opercularis; CP: cerebral peduncle; PrCG: precentral gyrus; SMG: supramarginal gyrus

3.4.3 Relationships Between HSFC model-related Connections and Speech Behaviors. Because not all behaviors survived permutation thresholding for the whole brain DTI analyses above, an exploratory analysis was conducted to correlate fiber counts between each behavior and the fiber counts where the AOS-Aphasia group demonstrated significantly fewer connections when compared to the Aphasia Only group. That is, fiber

counts from the fourteen regions presented in Table 3.5 were correlated with each behavior. Bonferroni corrected level of significance was set to 0.003 ($p=0.05/15$ comparisons=0.003). Because lesion size could influence the relationship between behavior and fiber counts, a second correlation analysis was conducted to determine the strength of these relationships when lesion volume was included as a co-factor. Correlation coefficients between behaviors that had at least one significant connection are presented in Table 3.9 below.

Notably, the following variables were not significantly correlated with any of the chosen connections when Bonferroni corrections were applied: peak frequency, peak amplitude, VOT-SD_{voiceless}, 2 Hz, 4 Hz, 16 Hz, and 32 Hz band energy (all $p>0.003$).

3.5 SUMMARY OF ALL RESULTS

The overall findings show that individuals with AOS performed differently than those without AOS on a number of measures. Importantly, the LDA results demonstrated that 16 Hz energy, VOT-SD_{voiced}, and the proportion of sound errors produced in a two-minute speech sample have the greatest weight in classification of AOS. Interestingly, although the nPVI-V coefficients differed significantly between groups, and have been previously reported to be important in differential diagnosis, this variable demonstrated less discriminative value when compared to other measures of speech production included in the LDA.

Table 3.9

Correlation coefficients for regional fiber counts and speech production errors

Listed here are behaviors that have at least one significant correlation with inter-regional fiber counts. *Indicates correlations that are significant at a Bonferroni corrected p -value of 0.003. **Indicates correlations that remain significant when corrected for lesion volume

	1 Hz Band	8 Hz Band	Syll/ Sec	WPM	Prod.	nPVI-V	Err (ph)	Err (D)	VOT- SD_{voice}
LPoCG-LaCR	-0.47*	0.52**	0.65**	0.66**	0.50**	0.27	-0.47*	-0.53*	-0.42*
LPrCG-LaCR	-0.44*	0.45*	0.64**	0.65**	0.58**	0.3	-0.50*	-0.49*	-0.53*
LPoCG-LAG	-0.45*	0.22	0.53*	0.48*	0.4	0.39	-0.49*	-0.40*	-0.07
RPrCG-LCaud	-0.18	0.23	0.35	0.321	0.2	0.34	-0.39	-0.39	-0.13
LSTG-LGP	-0.15	0.22	0.37	0.328	0.09	0.41*	-0.24	-0.3	-0.06
LPrCG-LIFGpo	-0.35	0.30	0.54*	0.54*	0.41*	0.45*	-0.52*	-0.53*	-0.34
LCaud-LpCR	-0.33	0.18	0.52*	0.50*	0.27	0.42*	-0.50*	-0.48*	-0.25
LIFGorb-LPoCG	-0.33	0.28	0.58*	0.60*	0.39	0.23	-0.54*	-0.49*	-0.32
LAG-LsCR	-0.48*	0.35	0.54*	0.50*	0.49*	0.22	-0.56*	-0.42*	-0.16
LCaud-LsCR	-0.41*	0.07	0.52*	0.48*	0.28	0.43*	-0.53*	-0.57*	-0.27*
LSMG-LsCR	-0.59*	0.58**	0.66**	0.64**	0.58**	0.44*	-0.59*	-0.53*	-0.43*
LPrCG-LSMG	-0.51*	0.28	0.61*	0.58*	0.32	0.41*	-0.54*	-0.49*	-0.28
LCaud-RsCR	-0.13	0.08	0.26	0.22	0.083	0.16	-0.32	-0.37*	-0.06
LAG-RIFGpo	-0.4	0.52*	0.53*	0.47*	0.49*	0.24	-0.54*	-0.46*	-0.11

With regard to the neuroanatomical representation of these variables, when controlling for aphasia severity, the variables previously reported to facilitate clinical classification of AOS (i.e., distortion errors, PVI) were predicted by damage to cortical areas associated with sensorimotor processes – the premotor, primary motor and sensory cortices, with additional involvement of underlying white matter (SLF) and the SMG predicting nPVI-V. Phonological errors, which occurred in both the Aphasia Only and AOS-Aphasia groups, were not predicted by damage to sensorimotor areas. Instead, these errors were predicted by damage to the AG, middle occipital gyrus (MOG) and the thalamic radiation. Furthermore, DTI results, both the whole brain and exploratory correlations, suggest that connections between these and other dorsal stream regions mediate specific behaviors crucial to articulate speech. The implications of these results will be discussed in detail in the next chapter.

CHAPTER 4

DISCUSSION

4.1 GENERAL DISCUSSION

Despite the number of studies that have endeavored to identify speech characteristics consistent with AOS, there is no universally adopted “gold standard” for its diagnosis. In clinical practice, many SLPs must rely on their perceptual evaluation of a patient’s speech, but perceptual evaluation is subject to the pitfalls of categorical perception (Buckingham & Yule, 1987; Code, 1998) and lack of inter-rater reliability (e.g., Haley et al., 2012; Mumby, Bowen, & Hesketh, 2007). The same problems can contaminate diagnostic labeling in research studies, as research groups often adopt their own diagnostic criteria, which may not be uniform across sites (McNeil et al., 1997; Mumby et al., 2007). As pervasively discussed in studies pertaining to AOS, these challenges are a burden to its clinical management.

Recent work has sought to decrease clinician bias in AOS diagnosis through the study of acoustic measures, which are inherently more objective (Ballard et al., 2016; Cunningham et al., 2015; Vergis et al., 2014). Accordingly, the purposes of this study were to 1) evaluate several measures (VOT, nPVI, EMS variables) and determine which had the greatest weight in diagnostic classification, 2) determine the perceptual correlates (i.e., ASRS items) of these objective measures, and 3) identify regional and network damage predictive of performance on these measures. This study is unique in that to our knowledge, no study has attempted to compare several objective measures within the

same sample of individuals with AOS. Furthermore, no study has yet to relate these behaviors to patterns of post-stroke cortical or white matter damage to regions implicated in contemporary models of speech production.

Overall, the results of this study show that individuals with AOS and concomitant aphasia can indeed be distinguished from those with aphasia only, and that objective measures of speech production are satisfactory for classification. In the sections to follow, these results will be discussed with regard to each of the specific aims of this study. A discussion of the theoretical implications of these findings will conclude this chapter.

4.2 SPECIFIC AIMS 1A AND 1B: IMPROVING THE ACCURACY OF DIAGNOSTIC CLASSIFICATION

The overall purpose of Aim 1 was to determine if objective measures can predict diagnostic classification, and if so, to identify the measures that account for the greatest variance in this prediction. Furthermore, this aim sought to determine the extent that those with AOS could be classified into subgroups based on the presence of distortion errors or lack of variability in rhythm (i.e., “equal and excess stress;” Kent & Rosenbek, 1983), arguably measured by nPVI-V (Vergis et al., 2014). In this section, the current results will be compared to those of recent studies (e.g., Ballard et al., 2016; Vergis et al., 2014), focusing on the clinical relevance of these collective findings. Within this discussion, the concept of AOS subtypes will be considered.

4.2.1 Comparisons Across Studies. Recently, Ballard et al. (2016) demonstrated that rPVI coefficients from words with a WS stress pattern (rPVI-WS) and scores from the increasing word-length subtest of the ABA-2 (Dabul, 2000; evaluates production

upon trials where words increase in length/complexity, i.e., *cat*, *catapult*, *catastrophe*) reliably classified those with AOS and aphasia from those with aphasia only (positive predictive value = 90.91%; negative predictive value = 87.18%). In the current study, the PVI-WS and PVI-SW coefficients obtained from the subset of participants with available PNTs did not reveal the same patterns of performance as reported in prior work (i.e., Ballard et al., 2016; Vergis et al., 2014). Both Ballard et al. (2016) and Vergis et al. (2014) showed that individuals with AOS and concomitant aphasia have significantly smaller rPVI coefficients when compared to individuals with aphasia only. In this study, there were no significant between-groups differences in word-level rPVI coefficients. Importantly, although the current AOS-Aphasia group had comparable rPVI coefficients when compared to those of Ballard et al. and Vergis et al., the Aphasia Only group demonstrated similarly reduced rPVI coefficients, with a large standard deviation (SD rPVI-WS=30.25).

There are several possibilities for these differences, as the current study differs from the aforementioned studies with regard to participant characteristics and the stimuli utilized to obtain rPVI coefficients. Vergis et al.'s (2014) study was small and included nine individuals with AOS *and* aphasia, eight with aphasia only, and eight age-matched controls. Ballard et al.'s (2016) sample was larger. It included 35 with AOS and concomitant aphasia and 37 with aphasia only, with seventeen of these datasets taken from the clinical groups reported by Vergis et al. (2014). Ballard et al.'s sample was indeed larger than the current sample, but mean WAB AQ scores did not differ when compared to the participants included here. However, it is difficult to compare further the

production abilities of the participants across studies as no other diagnostic measures overlapped between this study and Ballard et al.'s.

It is likely that discrepancies in word-level rPVI can be attributed to differences in articulatory demands for the stimuli used here. That is, some of the items obtained from the PNT had cluster onsets, but Ballard et al. and Vergis et al. utilized words with simpler CVCVCV structures. Production demands of the words used to obtain rPVI-WS and SW in this study were arguably greater, and this factor alone could have affected the duration of speech sound production across syllables (Munson, 2001). Additionally, although the number of tokens differed in this study compared to the others, Ballard et al. (2016) recommended use of at least five target words or multiple repetitions of a single word to obtain rPVI-WS, a recommendation met here.¹² Some participants in this study demonstrated multiple attempts at production, and all complete attempts were segmented and used for rPVI calculation. In cases where participants were not consistent across productions, this could have introduced more variability in production for both the Aphasia Only and AOS-Aphasia groups, yielding a null result for between-groups comparisons.

Nevertheless, despite the differences in word-level rPVI discussed above, the current results suggest that nPVI-V, obtained from connected speech samples, is indeed different in AOS compared to those with aphasia only. Therefore, to compare further the current results with those of Ballard and colleagues (2016), one final series of post-hoc analyses was conducted to compare 1) Ballard et al.'s predictors (word-level PVI, ABA-2 repeated trials subtest) to similar predictors obtained here (nPVI-V and the proportion of

¹² Full list of items can be found in Appendix C. The following words were used to obtain rPVI-WS: *thermometer*, *banana*, *binoculars*, *volcano*, and *piano*.

sound errors), and 2) the results of the current classification model that included the EMS variables to the model that was analogous to Ballard et al.'s (i.e., the model with nPVI-V and sound errors).

To this end, two separate logistic regression models were run – the first with nPVI-V and proportion of sound errors (essentially a replication of Ballard et al.), and the second with 4 and 16 Hz energy and the proportion of sound errors (the unique model obtained here). Results of each logistic regression analyses show that both models significantly accounted for a large proportion of variance¹³ (nPVI-V/sound error model: 74.5%; 4Hz/16Hz/sound error model: 83.4%), and have high *overall* accuracy in diagnostic classification (nPVI-V/sound error model: 86.2%; 4Hz/16Hz/sound error model: 94%). Importantly, the model that included the 4Hz and 16Hz measures predicted AOS classification with 90% accuracy, whereas the model that included nPVI-V was 80% accurate for classifying individuals with AOS. It should be noted that Ballard et al. (2016) used a different method of prediction, and prediction was based on a number variables in addition to measures of speech production (i.e., demographics, comprehension, naming, working memory, reading performance, auditory word discrimination, and results from an oral motor movement task). Nevertheless, the nPVI-V model from the current study corroborates Ballard et al.'s model with an independent and relatively large sample (here: N=59; Ballard et al.: N=72), while the 4Hz/16Hz model expands upon efforts to identify more objective means to distinguish individuals with AOS and concomitant aphasia from those with aphasia only.

¹³ Chi Square statistics for each logistic regression model are as follows: nPVI-V/Sound error model: $\chi^2(2)=45.12, p<0.001$; 4Hz/16Hz/Sound error model: $\chi^2(3)=53.09, p<0.001$

This is the first study, to our knowledge, that has used amplitude modulation spectra to classify a large sample of speakers with AOS from those without. The diagnostic value of these measures remains in rather early stages of investigation, but the results obtained here are promising. Future research investigating objective classification of AOS should consider implementing amplitude modulation measures, as replication of these results across clinical and research sites could have important, and practical, clinical applications. First, these variables had high accuracy in AOS classification, and variables representing low frequency modulation rates have also demonstrated similar accuracy in discriminating between the dysarthria subtypes (Liss et al., 2010). Second, although the use of PVI coefficients appears to be a viable means of differential diagnosis for AOS (e.g., Ballard et al., 2016), obtaining PVI coefficients can be time consuming and rather cumbersome. To obtain PVI, audio recordings from participant speech samples need to be imported into some speech analysis software (e.g., Praat; (Boersma & Weenink, 2001), and vocalic intervals must be segmented (often manually) so that durations can be obtained for PVI coefficient calculation. This process is not realistic in clinical practice.

Aside from these technical details, the extent that PVI indeed correlates with perceptual measures reflective of rhythm are questionable (Lowit, 2014), and others (Tilsen & Johnson, 2008) suggest that segmental duration measures ignore a great deal of important information about the acoustic signal within each C or V segment. Here, nPVI-V was significantly correlated with all but one measure on the ASRS, suggesting that it is sensitive to production behaviors that may also contribute to rhythmic impairment, but that it lacks specificity as a clinical marker unique to speech rhythm. On the other hand, band prominence at 4 Hz and 16 Hz was correlated with only a subset of ASRS items

(see Table 3.3; distorted substitutions, increased distortions with increased utterance length, inaccurate AMRs, sound distortions, lengthened intersegment durations, and audible/visible articulatory groping), meaning these acoustic variables may be more sensitive to studying specific aspects of production that are disrupted in AOS. Of course, further research is warranted regarding this claim.

Finally, a discussion of the clinical significance of these findings, pertaining to the use of the ASRS (Strand et al., 2014) as the “gold standard” for AOS classification, is warranted. As discussed thoroughly in Chapter 1, a widely used, well-validated and reliable psychometric assessment for the differential diagnosis of AOS does not exist. The one current assessment specific to AOS, the *ABA-2* (Dabul, 2000), is limited by the fact that it does not provide clinicians a method to determine if sound-level production errors are phonologic or articulatory in nature. The ASRS offers a preliminary solution to this problem. However, although the scale was developed by expert consensus of Mayo Clinic clinicians, and preliminary validation shows that the ASRS aligns well with expert diagnosis (Strand et al., 2014), it has not been validated in a multi-site study. It has also more frequently been used in the diagnosis of progressive AOS (e.g., Josephs et al., 2014; Josephs et al., 2012; Josephs et al., 2006; Strand et al., 2014; Whitwell, Duffy, Strand, Machulda, et al., 2013; Whitwell, Duffy, Strand, Xia, et al., 2013), and it is rather uncertain whether the behavioral characteristics of stroke-induced AOS and progressive AOS completely overlapping (Duffy & Josephs, 2012).

Accordingly, continued validation of the ASRS is warranted. Results from the current study can be useful in informing further study of the ASRS as a diagnostic tool, especially with regard to items that may or may not be pertinent to the differential

diagnosis of post-stroke AOS. For example, results from the LDA showed that the items most predictive of group classification (i.e., 4 and 16 Hz band prominence, VOT-SD_{voiced}) were correlated with a subset of ASRS items (Table 3.3). It is possible that this subset of items may be sufficient to aid in the diagnosis of AOS. A caveat here is that because the ASRS was used as the gold standard for the LDA, validation with another, unrelated measure (i.e., expert diagnosis) would be necessary to avoid circularity. Moreover, this study did not include a group of speakers with only dysarthria, a necessary comparison for further validation of the clinical utility of any measure intended for use to diagnose AOS.

4.3 SPECIFIC AIM 1C: EXPLORING THE EXISTENCE AOS SUBTYPES

One of the purposes of Aim 1 was to determine the extent that the participants with AOS could be classified into “subtypes” (Croot, 2002; Duffy & Josephs, 2012; Feiken & Jonkers, 2012). In rather small samples, the progressive AOS literature has provided preliminary evidence for the existence of AOS subtypes that can be characterized behaviorally and neuroanatomically (Duffy & Josephs, 2012; Josephs et al., 2013). However, more recent evidence suggests that subgroups are related to the presence of concomitant language difficulties in this population (Duffy et al., 2015; Duffy et al., 2013; Josephs et al., 2013).

nPVI-V coefficients obtained for participants in the AOS-Aphasia group yielded two “subgroups,” divided evenly into those with nPVI-V values within (n=10) and below (n=10) 1.5 standard deviations of the Aphasias Only group’s mean nPVI-V. However, these “subgroups” did not differ in the proportion of distortion errors produced, and results from the LDA did not indicate a linear distinction between these subgroups.

Furthermore, there was a significant negative correlation between nPVI-V and the proportion of sound errors ($r_s = -0.68$, $p < 0.001$), suggesting that as sound errors increased, nPVI-V also decreased (and vice versa). This relationship seems to indicate that prosodic adjustments could be a failed compensatory strategy for some individuals. For example, Maas and colleagues (2015) proposed that the AOS subtypes might actually reflect differences in feedback processing of speech. In this case, if a speaker with AOS has disrupted feedforward processing, that speaker may plan and program production by smaller units, allowing the feedback system time to self-monitor. In turn, this may result in fewer sound level errors, but stress patterns and speech rate could be affected (Kent, Kent, Weismer, & Duffy, 2000; Spencer & Slocumb, 2007). However, because the current study had no explicit measure of feedback/feedforward processing, whether or not different speech patterns reflect differences in internal monitoring can only be speculated from these results.

4.4 SPECIFIC AIM 2: BRAIN REGIONS AND CONNECTIONS THAT SUPPORT FLUENT SPEECH

The purpose of Aim 2 was to identify patterns of regional and network damage that predicted the dependent variables obtained here. It was hypothesized that the VLSM analyses would reveal differences in patterns of brain damage that relate to phonetic (VOT, narrow transcription, 16 Hz energy), phonemic (phonologic paraphasias), and syllabic/prosodic measures (nPVI-V, 1 Hz, 4 Hz, and 8 Hz band prominence). Specifically, it was hypothesized that *phonetic* errors would be related to patterns of damage to motor and sensorimotor areas, with decreased connectivity between a) motor areas and regions involved in sound level (phonemic) production (e.g., decreased

connection between inferior frontal areas and BA6/M1) and b) connections between cerebellar and motor areas (Hickok, 2012, 2014). In contrast, it was hypothesized that *phonemic* errors would correlate with damage to posterior temporoparietal areas (Buchsbaum et al., 2011; Ueno & Lambon Ralph, 2013), or anterior dorsal stream areas (Schwartz et al., 2012), with damage between these connections disrupted, and the PrCG relatively spared. Finally, it was hypothesized that measures of rhythm and prosody would be predicted by damage to inferior frontal and posterior sensorimotor regions, with disrupted connections between these regions and motor areas.

Results from the VLSM and connectome analyses confirmed many of the hypotheses stated above. First, damage to cortical sensorimotor areas, and connections between these areas and those implicated in phonologic processes (i.e., AG), was predictive of phonetic errors and variability in VOT for initial voiced stop consonants (VOT-SD_{voiced}). Variables reflective of speech rhythm (i.e., nPVI-V, 4 and 8 Hz band prominence) were also predicted by damage to sensorimotor areas, with additional involvement of the MFG (nPVI-V) and STG regions (4 Hz). Fiber count connections from the exploratory region of interest analysis further implicated the role of underlying white matter connections (i.e., L aCR, L sCR) to pre- and post-central regions (8 Hz energy, VOT-SD_{voiced}, distortion errors) as well as basal ganglia structures (nPVI-V: L caudate-L pCR; nPVI-V and VOT-SD_{voiced}: L caudate-L sCR).

These patterns of damage generally contrasted with those predictive of phonemic errors. Damage predictive of phonemic errors in the current study supports prior work that higher-level sound errors arise from damage to posterior temporoparietal areas (Lambon Ralph, Ehsan, Baker, & Rogers, 2012; Ueno & Lambon Ralph, 2013; Ueno,

Saito, Rogers, & Lambon Ralph, 2011), rather than anterior dorsal stream areas (e.g., Schwartz et al., 2012). Although the correlations between fiber counts from the region of interest analysis were similar for phonemic and distortion errors (see Table 3.9), results from the whole brain DTI analysis (Table 3.8) indicate that distortion errors were uniquely predicted by damage connecting the PrCG to the MFG and SMG. These findings show a clear neuroanatomical distinction between lower-level articulatory variables (i.e., phonetic errors) from those attributed to higher-level phonemic encoding (i.e., phonemic paraphasias). Even though the AOS-Aphasia group had significantly more phonemic errors than either of the other two groups, the neuroanatomical distinction between phonemic and phonetic behaviors illustrates that these processes are subserved by separate neuroanatomical substrates.

As hypothesized, damage to the basal ganglia itself was not predictive of any of the variables included in this study, but significant correlations in white matter connections between basal ganglia structures and cortical sensorimotor areas underscores the importance of corticostriatal connections in the pathophysiology of AOS. It is well established that subcortical and cerebellar structures are associated with timing and coordination of complex movements (e.g., speech production). As indicated in Table 3.9, several behaviors were correlated with fiber counts involving at least one connection to basal ganglia structures (i.e., left caudate and putamen) and surrounding white matter pathways (i.e., portions of the corona radiata). Inspection of proportional damage to the basal ganglia revealed that across groups, only a small proportion of overall lesion maps was located in the basal ganglia (less than 2% located in the caudate and putamen); however, large portions of these regions themselves were indeed affected, more so for the

AOS group. This is most apparent in the caudate, which has been implicated in overt speech production (Eickhoff, Heim, Zilles, & Amunts, 2009; Silbert, Honey, Simony, Poeppel, & Hasson, 2014), especially when production demands increase (Sörös et al., 2006). For the AOS-Aphasia group, 84% of the putamen was damaged, and 31% of the caudate. For the Aphasia Only group, 71% of the putamen was damaged, but only 1.2% of the caudate. In the Stroke Control group, both regions were over 90% spared. Notably, the AOS-Aphasia and Aphasia only groups did not significantly differ with regard to lesion volume. Taken together, while basal ganglia structures themselves did not emerge as regions most predictive of behavioral findings, timing and coordinative processes necessary for articulate speech were likely disrupted due to impairments in corticostriatal loops, caused by damage to the basal ganglia and surrounding white matter pathways.

Finally, with regard to the insula, peak amplitude was the only variable predicted by a large portion of damage to this region. It should be noted that this result was obtained from the univariate analysis that included peak amplitude only, as the multivariate analyses with aphasia severity and peak amplitude yielded less than 10 significant voxels. Larger lesions are associated with aphasia severity; therefore, it is possible that this result is an artifact of lesion size and could be explained by the fact that this area is a common site of damage in large left hemisphere middle cerebral artery strokes (Caviness et al., 2002; Finley et al., 2003; Hillis, 1989; Kodumuri et al., 2016). Nevertheless, this result may shed light on the role of the insula in speech production – a region that has been heavily debated for decades. (Basilakos et al., 2015; Dronkers, 1996; Richardson et al., 2012).

Because speech amplitude and prosodic contours are governed by respiratory forces (Fox, 2000), the relationship between peak amplitude and insula damage could also be explained by hypotheses that the insula is responsible for the coordination of breathing during speech production (Ackermann & Riecker, 2010; Fedorenko, Fillmore, Smith, Bonilha, & Fridriksson, 2015). Respiratory compromise itself is not a feature of AOS (Duffy, 2005), but in the progressive AOS population, Duffy (2006) found that 26% of their sample of 70 individuals with progressive AOS demonstrated reduced words per breath group despite relatively good sustained phonation. This finding may indicate a disruption in pre-utterance respiratory planning, or simply be attributed to the fact that individuals with AOS may only be able to plan shorter utterances prior to speech initiation (Duffy, 2006; Duffy & Josephs, 2012; Winkworth, Davis, Adams, & Ellis, 1995). Taken together, these findings warrant further study into the integrity of respiratory planning in speakers with AOS, and whether damage to the insula is responsible for this impairment. Doing so may provide further detail informing how respiratory processes are coordinated for speech production, and whether speech breathing is indeed affected in disorders not commonly associated with respiratory processes *per se*.

To summarize, lesion damage predictive of speech production deficits was associated with areas responsible for storing somato-phoneme targets (SMG/S1), as well as areas implicated in the creation of motor phoneme programs (pre-central gyrus regions BA6/M1). Lesion damage predictive of nPVI-V and VOT variability was heavily localized to post-central region (PoCG and SMG), but the primary areas predictive of distortion errors were the MFG, PrCG (together comprising 72% of the statistical map),

and with a small (8%) region localized in the post-central gyrus. Each behavior, however, was to some extent predicted by damage to both pre- and post-central regions.

Accordingly, apraxic impairment may result from a disruption in the SMG's ability to program articulator position information from the auditory template stored in posterior temporoparietal areas (Callan, Callan, Tajima, & Akahane-Yamada, 2006), or the coding of motor programs themselves (PrCG area; Hickok, 2014), especially as articulatory sequences increase in length and complexity (i.e., a role of the MFG; Bohland & Guenther, 2006). These findings suggest that AOS is not just related to processes that occur during phonetic encoding (i.e., programming syllable-sized targets from the auditory template) or the actual creation of motor programs, but that some aspects of both processes may be affected. The theoretical implications of these results will be discussed in greater detail in the next section.

4.5 THEORETICAL IMPLICATIONS

There is a lack of agreement regarding the production unit that guides speech planning (Bohland & Guenther, 2006). Research has suggested that speech planning processes occur at different levels – from the individual phoneme (Aichert & Ziegler, 2004; Ziegler, 2005, 2009), to the syllable and even word or phrase levels (Levelt et al., 1999; Levelt & Wheeldon, 1994; Varley & Whiteside, 2001; Varley et al., 1999). Others (e.g., MacNeilage, 1998; Shattuck-Hufnagel, 1979) propose a more complex relationship between the phoneme within prosodic and syllabic hierarchies (e.g., MacNeilage, 1998; Shattuck-Hufnagel, 1979). Accordingly, many theories exist regarding the level of impairment in AOS. Most recently, Maas et al. (2015) speculated that AOS may occur due to impairment at syllabic or phonemic levels. It can be implied that differences in the

extent that these levels are impaired may lead to different behavioral manifestations or be reflective of particular patterns of brain damage (Duffy & Josephs, 2012; Duffy et al., 2014; Duffy et al., 2013; McNeil et al., 2004).

As discussed throughout Chapter 1, there is no widely accepted, theoretical explanation for the speech production errors that occur in AOS, but contemporary models (i.e., DIVA, HSFC) offer some explanations. It was expected that the results obtained here could garner support from the HSFC model (Hickok, 2012, 2014), and this expectation was partially fulfilled. However, some of the results obtained here can also be explained by the DIVA model and its more recent update, the gradient order DIVA model (GODIVA; Bohland, Bullock, & Guenther, 2010).

As per the discussion in Section 3.4, results from the neuroimaging analyses showed that behaviors associated with AOS result from damage to regions responsible for the creation of somato-phoneme targets from auditory-syllable targets (somatosensory areas), the activation of articulatory motor plans (PrCG; Bohland et al., 2010; Long et al., 2016), and the sequencing and initiation of speech motor commands (supplementary motor area; (Bohland et al., 2010). The involvement of these cortical regions in the current findings fits well with both the DIVA and HSFC models. However, some of the behaviors measured here were related to the integrity of white matter connections between the basal ganglia (caudate and putamen) and cortical regions, implicating corticostriatal loops in articulate speech. Unlike the GODIVA model, the HSFC does not yet acknowledge the role of subcortical structures (i.e., the basal ganglia) or regional white matter connections in these processes. As such, the HSFC does not provide a complete explanation for the neuroanatomical findings from this study.

Moreover, because the GODIVA and HSFC models are relatively new, where one model lacks, the other may provide explanation. Accordingly, the remainder of this section is dedicated to a discussion of concepts needed to explain apraxic speech behaviors that should be given greater attention in contemporary models. This point will address production models widely, drawing examples from many, but ultimately, this chapter will conclude with suggestions for refining current models and possible directions for future studies. The specific focus of this discussion pertains to the potential role of timing and temporal coordination in AOS, processes that likely contribute to apraxic deficits, but that are not explicitly detailed in most models. The theme of this section draws from Levelt's (1989) statement that "there is no lack of theories, but there is a great need of convergence" (p. 452; also cited in Guenther, 1995).

Mounting evidence from the perception literature emphasizes the importance of temporal coordination in speech processing (Ghitza, 2011, 2013; Ghitza et al., 2012; Ghitza & Greenberg, 2009; Giraud et al., 2007; Giraud & Poeppel, 2012; Hyafil, Fontolan, Kabdebon, Gutkin, & Giraud, 2015). Although most of this work has focused on the role of temporal information in perception, production and perception are closely linked. For example, the purpose of speech is to produce a verbal message that can be comprehended by the listener. Speech motor commands are executed to fulfill an auditory goal, not only for the listener, but also for the purpose of feedback monitoring (Guenther et al., 2004; Hickok, 2012a, 2012b, 2014a, 2014b; Houde & Nagarajan, 2011; Perkell, 2012).

Across most models, the syllable is an important unit for both production and perception processes. The HSFC depicts a bidirectional flow between the motor syllable

and motor phoneme commands. Relatedly, the GODIVA model simulates production via “cotemporal” activation of phonetic plans and syllabic frames. The architectures of these models illustrate that lower-level phonetic-motoric processes are intertwined with those occurring at the syllable level. The importance of a syllable-sized coordinative unit can be exemplified further by findings that endogenous neural oscillations occur at rates consistent with syllable production (i.e., theta range, 4-10 Hz), and that these theta oscillations may serve to entrain beta (15-30 Hz) and gamma (31-50 Hz) oscillations to facilitate the encoding of phonetic detail (i.e., the “fine structure”) of auditorily presented speech (Ghitza, 2011, 2013; Ghitza et al., 2012; Giraud & Poeppel, 2012). Drawing upon a large body of prior literature documenting the importance of the syllable in speech production (Bohland & Guenther, 2006; Ghitza, 2013; Ghitza et al., 2012; Guenther, 1995; Kotz & Schwarze, 2010; Levelt, 1999; Levelt et al., 1999; Levelt & Wheeldon, 1994; MacNeilage, 1998; Shattuck-Hufnagel, 2015; Ziegler, 2009), the results obtained here point towards a rhythmic-gestural (Tilsen & Johnson, 2008) interaction between motor programs at the level of the syllable, cascading to smaller (i.e., gestural) units. Notably, the rhythmic nature of speech is often attributed to syllabic architecture (typically with regard to the vowel), but pointed out by Kotz and Schwarze, (2010) many models do not often mention the importance of temporal information in speech production processes.

It is worth highlighting that three of the measures obtained in this study are temporal in nature. VOT is the duration between plosive release and the onset of voicing (a measure of inter-articulatory timing and coordination; Baum et al., 1990; Blumstein et al., 1980; Schirmer, 2004; Seddoh et al., 1996), PVI is derived by the variability of

successive vocalic durations, and the variables obtained from the envelope modulation spectra reflect amplitude modulation rates associated with syllabic production, prosodic contours, and phoneme durations. Pervasive in the AOS literature is the hypothesis that speech planning and programming processes lack coordination (e.g., Buchwald, 2014; Itoh & Sasanuma, 1984; Kent & McNeil, 1987; Maas et al., 2008; McNeil et al., 2000; McNeil & Kent, 1990; Ziegler, 2009; Ziegler & von Cramon, 1986b). Timing is crucial for planning/programming any gross or fine motor movement (Kotz & Schwartz, 2010), and this is especially true for speech production, as articulate speech relies on both proper sequencing *and* timing of motor plans (Turk & Shattuck-Hufnagel, 2014). Work by Maas and colleagues (2008) suggested that the process of “structuring” motor plans is disrupted in AOS. According to this view, individuals with AOS have difficulties integrating the details of a motor program, and these details often include parameters for movement timing (Klapp, 2003). In fact, Ziegler and von Cramon have proposed that behavioral characteristics of AOS can be attributed to mistiming of speech processes (Ziegler & von Cramon, 1986b). Maas and colleagues’ hypothesis concurs with other studies suggesting that distortion errors are related to mistimed programs (Baum et al., 1990; Blumstein et al., 1980; Buchwald, 2014), and that these types of phonetic errors are indeed distinguishable from higher-level phoneme selection errors (Buchwald, 2014). However, many of these results obtained from individuals with AOS have not been studied beyond a string of syllables or short phrases. Therefore, it is uncertain how individuals with AOS keep up with the demands of ongoing, conversational speech when both linguistic (i.e., aphasia) *and* motor planning impairments may affect production processes. For this reason, the current study analyzed production measures from connected speech samples.

Results from the EMS variables suggest a relationship between the theta (syllabic, 4 Hz) and beta (phonetic, 16 Hz) rates during production. First, as evident in Figure 3.12, the trend between band prominence in the 4, 8 and 16 Hz bands was parallel across groups – energy within these modulation rates decreased between the Stroke Control and Aphasia Only group, and again between the Aphasia Only and AOS-Aphasia groups. Second, as band prominence in these frequency bands decreased across groups, 1 Hz energy increased, resulting in production characterized by a higher band prominence at 1 Hz in the AOS-Aphasia group. It may be that these slower frequency bands reflect some aspect of temporal control over speech production, and when temporal control was compromised, processes requiring temporal gating (i.e., phonetic production, intonation contours) were also affected (Ghitza & Greenberg, 2009 and Ghitza, 2011), as indicated by the correlations with sound-level errors (Figure 3.12). Relatedly, in an amplitude envelope modulation spectrum analysis using approximately 300,000 words obtained from 40 adults during conversational speech, Tilsen and Johnson (2008) showed an inverse relationship between speech disfluencies and 1 Hz rhythm (Tilsen & Johnson, 2008). Tilsen and Johnson (2008) explain this relationship as the result of vowel preservation due to decreased speech rate, but acknowledge the need for the development of a theoretical explanation for the relationship between speech rhythm and gestural production.

Before proceeding with further discussion, it should be addressed that the relationship between band prominence, sound level errors, and AOS severity could be epiphenomenal and simply reflect the fact that individuals with more severe AOS are more likely to have both a higher number of distortion errors and speech characterized

more so by atypical rhythm, perhaps to implement real-time feedback compensation. However, it seems more than coincidental that this pattern is very similar to the results obtained by Ghitza and Greenberg (2009) in a study that investigated speech intelligibility (measured by recognition errors) for compressed speech as a function of syllabic rhythm (manipulated by inserting silent pauses in the compressed speech signal). Ghitza and Greenberg (2009) showed that when participants were presented highly compressed speech, intelligibility was modulated by syllabic rhythm. Pause durations that deviated either rightward or leftward of 20-160 ms pause durations were associated with more errors. The authors interpreted these findings to suggest that the acoustic waveforms of the speech signals modulated by these pause durations are analogous to the theta range, and endogenous neural oscillations at this frequency range facilitated speech processing.

Ghitza stated that no "purely auditory or articulatory" (Ghitza, 2011, p. 4) model can yet account for the findings of their (Ghitza and Greenberg, 2009) study, but the results obtained here seem to suggest that articulatory processes may also be governed by similar perceptual constraints for speech processing. Here, as relative temporal dominance shifted away from the 4-8Hz range in either direction, the correlation between production errors increased, with higher error rates associated with both the 1 Hz and 16-32 Hz bands. Essentially, the results obtained here and those by Ghitza and Greenberg (2009) reveal a U-shaped relationship between temporal envelope modulation and errors – both for the domain of production (current study) and perception (Ghitza & Greenberg, 2009). Perhaps the current findings reflect that there is an optimal (syllabic-sized) modulation rate for speech monitoring; when speech motor processes cannot meet these

demands, monitoring is compromised, resulting in higher error rates. Further discussion on this point will be taken up in the following paragraphs.

For a model to account for AOS, it should explain how disruptions in temporal coordination at one level affect processes at another level. Although the current results regarding the relationship between rhythmic and segmental production seem to support this prediction, contemporary models do not completely account for how the syllable can serve as a rhythmic “pacer” for production. That syllable-sized units are important for phonetic-articulatory planning is by no means a novel idea in the AOS literature (see discussion above). The syllable, regarded as the level at which impaired planning/programming processes manifest, (Maas et al., 2008; Mailend & Maas, 2013; Varley & Whiteside, 2001; Varley et al., 1999), has been credited with a facilitatory role in phonetic planning (Ziegler, 2009). For example, following a retrospective analysis of 40 participants with AOS, Ziegler found that even when word length increased, the probability of production accuracy on words with disyllabic or trochaic feet was similar to that of monosyllabic tokens (Ziegler, 2009). Ziegler likened the role of syllabic structure as a “phonetic molecule with strong atomic bonds” (Ziegler, 2009, p. 657), emphasizing that the syllable provides some organizational structure to guide lower-level production processes.

Currently, some models account for the role of timing in speech production (e.g., DIVA: Guenther, 1995; GODIVA: Guenther et al., 2010; and “prosody first models”, Keating & Shattuck-Hufnagel, 2002; Shattuck-Hufnagel, 2015). In a review of studies that investigated various aspects of timing during speech production, Turk and Shattuck-Hufnagel (2014) acknowledged that the exact role of timing processes is largely

unknown; however, the authors suggest that timing may likely support feedback-controlled error correction. To speculate, because speech production relies upon intact monitoring processes, the maintenance of syllable-gesture relationships may facilitate temporal coordinative processes utilized by feedback processing. If production, like perception, is mediated by a “theta oscillator” (Ghitza, 2013), preservation of a slow-rate temporal structure may reflect a compensatory reliance on prosodic frames to facilitate planning (see also Ziegler, 2009; Keating & Shattuck-Hufnagel, 2002; Shattuck-Hufnagel, 2015; Ziegler, 2009). Relevant to the HSFC, Kotze and Schwarze (2010) suggest that temporal “when” information can guide processing of the “what” information, which aligns well with the HSFC's dual stream architecture. That is, the authors point out that acoustic information is mapped to articulatory/phonemic information in the dorsal stream, and the temporal information obtained from acoustic signals can guide further feedback processing. Several models (e.g., DIVA, SFC, HSFC) propose that the goal of speech production is to plan, program and execute motor commands so that when speech is produced, it matches a previously learned expectation. In both the HSFC and DIVA, syllable-sized auditory targets are the foundation for these goals, but unique to the HSFC is the proposal that auditory feedback monitors syllable-sized units (but see Guenther, 2014 and Niziolek and Guenther, 2013 for evidence that lower-level units can be guided by auditory feedback). Perhaps this temporal “when” information supports auditory feedback monitoring processes, which may be heavily relied upon in speakers with AOS (Jacks, 2008; Maas et al., 2015). For example, differences in band prominence across modulation rates for the AOS group may be reflective of compensatory mechanisms necessary for effective feedback control over

speech. Kent and Rosenbek (1983) suggested when speakers rely on feedback processes due to an impairment in feedforward control, speech production becomes slowed and with little variability in stress (i.e., “excess and equal stress”) because such pattern facilitates auditory monitoring.

With regard to the neuroanatomical correlates of these processes, the DIVA model posits that the “Auditory Error Map” is located within bilateral regions of the perisylvian cortex (Bohland & Guenther, 2006). In the current study, damage to these aforementioned areas implicated in auditory feedback monitoring was related to 4 Hz band prominence, suggesting that regions implicated in temporal modulations at this rate are also implicated in feedback monitoring. Behaviorally, there was a tradeoff between 4 Hz and 1 Hz band prominence, and interestingly, 1 Hz band prominence was predicted by integrity of right hemisphere connections between the R ACC and R pole of the STG, the R middle occipital gyrus (MOG) and L cerebellum, and the R MOG and L lingual gyrus (Table 3.8). These findings suggest that the surface timing of speech production was altered to preserve some semi-periodic production pattern (Turk & Shattuck-Hufnagel, 2013), mediated by right hemisphere white matter tracts (Poeppel, 2003).

The above explanations are not necessarily mutually exclusive, as it is generally accepted that error monitoring and speech timing are critically linked (Bohland & Guenther, 2006; Guenther, 1995, 2006; Houde & Nagarajan, 2011; Kotz & Schwartze, 2010; Perkell, 2012; Shattuck-Hufnagel, 2015). Further research is warranted to determine what mechanism is driving the behavioral findings obtained here, and whether or not study of the temporal nature of speech production can further inform the nature of apraxic deficits. Approaching production from this point of view makes sense from the

standpoint that treatments that involve rate/rhythm control show some success in treating non-fluent aphasia and AOS. A systematic review of the treatment literature (Ballard et al., 2015) showed that approaches that incorporate elements of rate/rhythm are beneficial for individuals with AOS (e.g., Albert, Sparks, & Helm, 1973; Belin et al., 1996; Boucher, Garcia, Fleurant, & Paradis, 2001; Brendel & Ziegler, 2008; Fujii & Wan, 2014; Norton, Zipse, Marchina, & Schlaug, 2009; Racette, Bard, & Peretz, 2006; Schlaug, Marchina, & Norton, 2008; Wambaugh et al., 2012). Some have suggested that individuals respond positively to such treatments due to compensation from the right hemisphere (Norton et al., 2009; Schlaug et al., 2008) or invoke the involvement of subcortical structures implicated in timing and coordination (i.e., basal ganglia: Stahl et al., 2011; for a review, see Fujii & Wan, 2014). Others have suggested that rate/rhythm controlled treatments may facilitate feedback control (Brendel & Ziegler, 2008; Fujii & Wan, 2014).

4.6 FINAL REMARKS

This study is the first, to our knowledge, to provide a comprehensive investigation of phonetic-level behaviors, and their neuroanatomical correlates, in a large sample of post-stroke individuals. Moreover, in light of recent attempts to obtain objective measures to classify individuals with AOS from those without, this was the first study to compare a several objective variables to determine which measure has the greatest discriminative weight in diagnostic classification. Aside from informing clinical practice, these results can be used to inform further contemporary models of speech production. Importantly, these findings highlight the importance of temporal coordination in speech production, how this can be perturbed in AOS, and the cortical, subcortical and white

matter connections that, when damaged, mediate this disruption. Future study should consider the extent that the results obtained here are reflective of disruptions in phonetic encoding itself, feedforward/feedback control of speech production, or some combination of both of these processes.

Importantly, growing evidence from patient populations and normal individuals suggests that cortical motor and sensorimotor regions are responsible for different aspects of speech planning and programming, and that different error types can be localized neuroanatomically (e.g., initiation and timing, sound sequencing, handling increased production demands, creating motor plans, among others; Blumstein & Baum, 2016; Bohland et al., 2010; Eickhoff et al., 2009; Hartwigsen et al., 2013; Moser et al., 2009; Price, 2012). No model has yet to provide a fine-grained treatment of these regions throughout the production process. As such, the results obtained here can be used to further refine existing models by 1) providing information regarding production impairments attributed to lower-level articulatory processes (planning and programming), and 2) identifying damage to both regional and network connections that support these specific processes. Taken together, this study provides a more fine-grained account of details regarding speech planning and programming required for production, the neuroanatomical regions that support these processes, and a discussion to guide future research.

REFERENCES

- Ackermann, H., & Riecker, A. (2010). The contribution (s) of the insula to speech production: a review of the clinical and functional imaging literature. *Brain Structure and Function*, 214(5-6), 419-433.
- Aichert, I., & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain Lang*, 88(1), 148-159.
- Albert, M. L., Sparks, R. W., & Helm, N. A. (1973). Melodic intonation therapy for aphasia. *Arch Neurol*, 29(2), 130-131.
- Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*, 40(3), 351-373.
- Ashburner, J., Barnes, G., Chen, C., Daunizeau, J., Flandin, G., Friston, K., . . . Litvak, V. (2012). SPM8 manual. *Functional Imaging Laboratory, Institute of Neurology*.
- Auzou, P., Ozsancak, C., Morris, R. J., Jan, M., Eustache, F., & Hannequin, D. (2000). Voice onset time in aphasia, apraxia of speech and dysarthria: a review. *Clinical Linguistics & Phonetics*, 14(2), 131-150.
- Baldo, J. V., Wilkins, D. P., Ogar, J., Willock, S., & Dronkers, N. (2011). Role of the precentral gyrus of the insula in complex articulation. *Cortex*, 47(7), 800-807. doi:10.1016/j.cortex.2010.07.001
- Ballard, K. J., Azizi, L., Duffy, J. R., McNeil, M. R., Halaki, M., O'Dwyer, N., . . . Robin, D. A. (2016). A predictive model for diagnosing stroke-related apraxia of speech. *Neuropsychologia*, 81, 129-139.
- Ballard, K. J., Barlow, J. A., & Robin, D. A. (2001). The underlying nature of apraxia of speech: A critical evaluation of Varley and Whiteside's dual route speech encoding hypothesis. *Aphasiology*, 15(1), 50-57.
- Ballard, K. J., Granier, J. P., & Robin, D. A. (2000). Understanding the nature of apraxia of speech: Theory, analysis and treatment. *Aphasiology*, 14(10), 969-995.
- Ballard, K. J., Tourville, J. A., & Robin, D. A. (2014). Behavioral, computational, and neuroimaging studies of acquired apraxia of speech. *Front Hum Neurosci*, 8, 892. doi:10.3389/fnhum.2014.00892
- Ballard, K. J., Wambaugh, J. L., Duffy, J. R., Layfield, C., Maas, E., Mauszycki, S., & McNeil, M. R. (2015). Treatment for Acquired Apraxia of Speech: A Systematic Review of Intervention Research between 2004 and 2012. *American Journal of Speech-Language Pathology*. doi:10.1044/2015_AJSLP-14-0118
- Basilakos, A., Fillmore, P. T., Rorden, C., Guo, D., Bonilha, L., & Fridriksson, J. (2014). Regional white matter damage predicts speech fluency in chronic post-stroke aphasia. *Front Hum Neurosci*, 8, 845. doi:10.3389/fnhum.2014.00845
- Basilakos, A., Rorden, C., Bonilha, L., Moser, D., & Fridriksson, J. (2015). Patterns of Poststroke Brain Damage That Predict Speech Production Errors in Apraxia of Speech and Aphasia Dissociate. *Stroke*. doi:10.1161/STROKEAHA.115.009211

- Baum, S. R., Blumstein, S. E., Naeser, M. A., & Palumbo, C. L. (1990). Temporal dimensions of consonant and vowel production: An acoustic and CT scan analysis of aphasic speech. *Brain and language*, 39(1), 33-56.
- Behroozmand, R., Karvelis, L., Liu, H., & Larson, C. R. (2009). Vocalization-induced enhancement of the auditory cortex responsiveness during voice F 0 feedback perturbation. *Clinical Neurophysiology*, 120(7), 1303-1312.
- Belin, P., Zilbovicius, M., Remy, P., Francois, C., Guillaume, S., Chain, F., . . . Samson, Y. (1996). Recovery from nonfluent aphasia after melodic intonation therapy A PET study. *Neurology*, 47(6), 1504-1511.
- Blumstein, S. E., Alexander, M. P., Ryalls, J. H., Katz, W., & Dworetzky, B. (1987). On the nature of the foreign accent syndrome: A case study. *Brain and language*, 31(2), 215-244.
- Blumstein, S. E., Baker, E., & Goodglass, H. (1977). Phonological factors in auditory comprehension in aphasia. *Neuropsychologia*, 15(1), 19-30.
- Blumstein, S. E., & Baum, S. R. (2016). Neurobiology of speech production: Perspective from neuropsychology and neurolinguistics. In G. Hickok & S. L. Small (Eds.), *Neurobiology of Language* (pp. 689-697). Boston, MA: Elsevier.
- Blumstein, S. E., Cooper, W. E., Goodglass, H., Statlender, S., & Gottlieb, J. (1980). Production deficits in aphasia: A voice-onset time analysis. *Brain and language*, 9(2), 153-170.
- Boersma, P., & Weenink, D. (2001). Praat, a system for doing phonetics by computer.
- Bohland, J. W., Bullock, D., & Guenther, F. H. (2010). Neural representations and mechanisms for the performance of simple speech sequences. *Journal of Cognitive Neuroscience*, 22(7), 1504-1529.
- Bohland, J. W., & Guenther, F. H. (2006). An fMRI investigation of syllable sequence production. *Neuroimage*, 32(2), 821-841.
- Bonilha, L., Nesland, T., Rorden, C., Fillmore, P., Ratnayake, R. P., & Fridriksson, J. (2014). Mapping remote subcortical ramifications of injury after ischemic strokes. *Behav Neurol*, 2014, 215380. doi:10.1155/2014/215380
- Bonilha, L., Rorden, C., & Fridriksson, J. (2014). Assessing the clinical effect of residual cortical disconnection after ischemic strokes. *Stroke*, 45(4), 988-993. doi:10.1161/STROKEAHA.113.004137
- Boucher, V., Garcia, L., Fleurant, J., & Paradis, J. (2001). Variable efficacy of rhythm and tone in melody-based interventions: implications for the assumption of a right-hemisphere facilitation in non-fluent aphasia. *Aphasiology*, 15(2), 131-149.
- Brady, M. C., Kelly, H., Godwin, J., & Enderby, P. (2014). Speech and language therapy for aphasia following stroke. *Cochrane Database Syst Rev*, 5(CD000425). doi:10.1002/14651858.CD000425.pub2
- Brendel, B., & Ziegler, W. (2008). Effectiveness of metrical pacing in the treatment of apraxia of speech. *Aphasiology*, 22(1), 77-102.
- Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D'Esposito, M., & Hickok, G. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory - an aggregate analysis of lesion and fMRI data. *Brain Lang*, 119(3), 119-128. doi:10.1016/j.bandl.2010.12.001

- Buchwald, A. (2014). Phonetic Processing. In M. A. Goldrick, V. Ferreira, & M. Miozzo (Eds.), *The Oxford handbook of language production* (pp. 245-258). New York, NY: Oxford University Press.
- Buckingham, H. W., & Yule, G. (1987). Phonemic false evaluation: Theoretical and clinical aspects. *Clinical Linguistics & Phonetics*, 1(2), 113-125.
- Callan, A. M., Callan, D. E., Tajima, K., & Akahane-Yamada, R. (2006). Neural processes involved with perception of non-native durational contrasts. *Neuroreport*, 17(12), 1353-1357.
- Carbonell, K. M., Lester, R. A., Story, B. H., & Lotto, A. J. (2015). Discriminating simulated vocal tremor source using amplitude modulation spectra. *Journal of Voice*, 29(2), 140-147.
- Catani, M., Dell'Acqua, F., Vergani, F., Malik, F., Hodge, H., Roy, P., . . . Thiebaut de Schotten, M. (2012). Short frontal lobe connections of the human brain. *Cortex*, 48(2), 273-291.
- Catani, M., & Mesulam, M. (2008). The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. *Cortex*, 44(8), 953-961. doi:10.1016/j.cortex.2008.04.002
- Catani, M., Mesulam, M. M., Jakobsen, E., Malik, F., Martersteck, A., Wieneke, C., . . . Weintraub, S. (2013). A novel frontal pathway underlies verbal fluency in primary progressive aphasia. *Brain*, 136(8), 2619-2628.
- Caviness, V., Makris, N., Montinaro, E., Sahin, N., Bates, J., Schwamm, L., . . . Kennedy, D. (2002). Anatomy of Stroke, Part I An MRI-Based Topographic and Volumetric System of Analysis. *Stroke*, 33(11), 2549-2556.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284.
- Code, C. (1998). Models, theories and heuristics in apraxia of speech. *Clinical Linguistics & Phonetics*, 12(1), 47-65.
- Cole, M. F., & Cole, M. E. (1971). *Pierre Marie's papers on speech disorders*. New York, NY: Hafner.
- Croot, K. (2002). *Diagnosis of AOS: definition and criteria*. Paper presented at the Seminars in Speech and Language.
- Crouzet, O., & Ainsworth, W. A. (2001). *On the various influences of envelope information on the perception of speech in adverse conditions: An analysis of between-channel envelope correlation*. Paper presented at the Workshop on Consistent and Reliable Acoustic Cues for Sound Analysis, Aalborg, Denmark.
- Cunningham, K. T., Haley, K. L., & Jacks, A. (2015). Speech sound distortions in aphasia and apraxia of speech: reliability and diagnostic significance. *Aphasiology*(ahead-of-print), 1-18.
- Dabul, B. L. (2000). *Apraxia battery for adults*. Dallas, TX: Pro-Ed.
- Dronkers, N. (1996). A new brain region for coordinating speech articulation. *Nature*, 384(6605), 159-161. doi:10.1038/384159a0
- Dronkers, N., & Ogar, J. (2004). Brain areas involved in speech production. *Brain*, 127, 1461-1462. doi:DOI 10.1093/brain/awh233

- Dronkers, N., Ogar, J., Willock, S., & Wilkins, D. P. (2004). Confirming the role of the insula in coordinating complex but not simple articulatory movements. *Brain and language*, 91(1), 23-24. doi:DOI 10.1016/j.bandl.2004.06.016
- Duffy, J. (2005). *Motor speech disorders: Substrates, differential diagnosis and management*. . St. Louis, M.O. : Mosby.
- Duffy, J. R. (1998). Stroke with dysarthria: evaluate and treat; garden variety or down the garden path? *Semin Speech Lang*, 19(1), 93-98; quiz 99. doi:10.1055/s-2008-1064039
- Duffy, J. R. (2006). Apraxia of speech in degenerative neurologic disease. *Aphasiology*, 20(6), 511-527.
- Duffy, J. R., & Josephs, K. A. (2012). The diagnosis and understanding of apraxia of speech: why including neurodegenerative etiologies may be important. *J Speech Lang Hear Res*, 55(5), S1518-S1522. doi:10.1044/1092-4388(2012/11-0309)
- Duffy, J. R., & Josephs, K. A. (2012). The diagnosis and understanding of apraxia of speech: Why including neurodegenerative etiologies may be important. *Journal of Speech, Language, and Hearing Research*, 55(5), S1518-S1522.
- Duffy, J. R., Peach, R. K., & Strand, E. A. (2007). Progressive apraxia of speech as a sign of motor neuron disease. *Am J Speech Lang Pathol*, 16(3), 198-208. doi:10.1044/1058-0360(2007/025)
- Duffy, J. R., Strand, E. A., Clark, H., Machulda, M., Whitwell, J. L., & Josephs, K. A. (2015). Primary Progressive Apraxia of Speech: Clinical Features and Acoustic and Neurologic Correlates. *Am J Speech Lang Pathol*. doi:10.1044/2015_AJSLP-14-0174
- Duffy, J. R., Strand, E. A., Clark, H., Machulda, M., Whitwell, J. L., & Josephs, K. A. (2015). Primary Progressive Apraxia of Speech: Clinical Features and Acoustic and Neurologic Correlates. *American Journal of Speech-Language Pathology*, 24(2), 88-100.
- Duffy, J. R., Strand, E. A., & Josephs, K. A. (2014). Motor Speech Disorders Associated with Primary Progressive Aphasia. *Aphasiology*, 28(8-9), 1004-1017. doi:10.1080/02687038.2013.869307
- Duffy, J. R., Strand, E. E., Whitwell, J. L., & Josephs, K. A. (2013). *Progressive Apraxia of Speech: Might There Be Subtypes?* Paper presented at the Clinical Aphasiology Conference, Tucson, AZ.
- Eickhoff, S. B., Heim, S., Zilles, K., & Amunts, K. (2009). A systems perspective on the effective connectivity of overt speech production. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 367(1896), 2399-2421.
- Enderby, P. (1980). Frenchay dysarthria assessment. *British Journal of Disorders of Communication*, 15(3), 165-173.
- Faria, A. V., Joel, S. E., Zhang, Y., Oishi, K., van Zijl, P. C., Miller, M. I., . . . Mori, S. (2012). Atlas-based analysis of resting-state functional connectivity: evaluation for reproducibility and multi-modal anatomy-function correlation studies. *Neuroimage*, 61(3), 613-621. doi:10.1016/j.neuroimage.2012.03.078
- Fedorenko, E., Fillmore, P., Smith, K., Bonilha, L., & Fridriksson, J. (2015). The superior precentral gyrus of the insula does not appear to be functionally specialized for articulation. *J Neurophysiol*, 113(7), 2376-2382. doi:10.1152/jn.00214.2014

- Feiken, J. F., & Jonkers, R. (2012). *A division into three subtypes of apraxia of speech: ideomotor, kinetic and ideational*. Paper presented at the 42nd Clinical Aphasiology Conference, Lake Tahoe, CA.
<http://aphasiology.pitt.edu/archive/00002349/>
- Finley, A., Saver, J., Alger, J., Prgenzer, M., Leary, M., Ovbiagele, B., . . . Kidwell, C. (2003). *Diffusion weighted imaging assessment of insular vulnerability in acute middle cerebral artery infarctions*. Paper presented at the Stroke.
- Fox, A. (2000). *Prosodic features and prosodic structure: The phonology of suprasegmentals*: Oxford University Press.
- Freeman, F. J., Sands, E. S., & Harris, K. S. (1978). Temporal coordination of phonation and articulation in a case of verbal apraxia: A voice onset time study. *Brain and Language*, 6(1), 106-111.
- Fridriksson, J., Guo, D., Fillmore, P., Holland, A., & Rorden, C. (2013). Damage to the anterior arcuate fasciculus predicts non-fluent speech production in aphasia. *Brain*, 136(Pt 11), 3451-3460. doi:10.1093/brain/awt267
- Fujii, S., & Wan, C. Y. (2014). The role of rhythm in speech and language rehabilitation: the SEP hypothesis.
- Galluzzi, C., Bureca, I., Guariglia, C., & Romani, C. (2015). Phonological simplifications, apraxia of speech and the interaction between phonological and phonetic processing. *Neuropsychologia*, 71, 64-83.
- Ghitza, O. (2011). Linking speech perception and neurophysiology: speech decoding guided by cascaded oscillators locked to the input rhythm. *Front Psychol*, 2.
- Ghitza, O. (2013). The theta-syllable: a unit of speech information defined by cortical function. *Front Psychol*, 4.
- Ghitza, O., Giraud, A.-L., & Poeppel, D. (2012). Neuronal oscillations and speech perception: critical-band temporal envelopes are the essence. *Frontiers in Human Neuroscience*, 6.
- Ghitza, O., & Greenberg, S. (2009). On the possible role of brain rhythms in speech perception: intelligibility of time-compressed speech with periodic and aperiodic insertions of silence. *Phonetica*, 66(1-2), 113-126.
- Giraud, A.-L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S., & Laufs, H. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron*, 56(6), 1127-1134.
- Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511-517.
- Golfopoulos, E., Tourville, J. A., Bohland, J. W., Ghosh, S. S., Nieto-Castanon, A., & Guenther, F. H. (2011). fMRI investigation of unexpected somatosensory feedback perturbation during speech. *Neuroimage*, 55(3), 1324-1338.
doi:10.1016/j.neuroimage.2010.12.065
- Golfopoulos, E., Tourville, J. A., & Guenther, F. H. (2010). The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *Neuroimage*, 52(3), 862-874. doi:10.1016/j.neuroimage.2009.10.023
- Goodglass, H. (1992). Diagnosis of conduction aphasia. *Conduction aphasia*, 39-49.
- Goodglass, H., Kaplan, E., & Barresi, B. (2000). *Boston Diagnostic Aphasia Examination-Third Edition*. San Antonio, TX: Pearson.

- Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in laboratory phonology*, 7(515-546).
- Graff-Radford, J., Jones, D. T., Strand, E. A., Rabinstein, A. A., Duffy, J. R., & Josephs, K. A. (2014). The neuroanatomy of pure apraxia of speech in stroke. *Brain Lang*, 129, 43-46. doi:10.1016/j.bandl.2014.01.004
- Greenlee, J. D., Behroozmand, R., Larson, C. R., Jackson, A. W., Chen, F., Hansen, D. R., . . . Howard, M. A. (2013). Sensory-motor interactions for vocal pitch monitoring in non-primary human auditory cortex. *PLoS One*, 8(4), e60783.
- Guenther, F. H. (1994). A neural network model of speech acquisition and motor equivalent speech production. *Biol Cybern*, 72(1), 43-53.
- Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological review*, 102(3), 594.
- Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychol Rev*, 102(3), 594-621.
- Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *J Commun Disord*, 39(5), 350-365.
- Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language*, 96(3), 280-301.
- Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychol Rev*, 105(4), 611-633.
- Guenther, F. H., Perkell, J. S., Maassen, B., Kent, R., Peters, H., van Lieshout, P., & Hulstijn, W. (2004). A neural model of speech production and its application to studies of the role of auditory feedback in speech. *Speech motor control in normal and disordered speech*, 29-49.
- Guenther, F. H., & Vladusich, T. (2012). A Neural Theory of Speech Acquisition and Production. *J Neurolinguistics*, 25(5), 408-422. doi:10.1016/j.jneuroling.2009.08.006
- Hagmann, P., Cammoun, L., Gigandet, X., Gerhard, S., Grant, P. E., Wedeen, V., . . . Sporns, O. (2010). MR connectomics: principles and challenges. *Journal of Neuroscience Methods*, 194(1), 34-45.
- Haley, K. L., Jacks, A., & Cunningham, K. T. (2013). Error variability and the differentiation between apraxia of speech and aphasia with phonemic paraphasia. *J Speech Lang Hear Res*, 56(3), 891-905. doi:10.1044/1092-4388(2012/12-0161)
- Haley, K. L., Jacks, A., de Riesthal, M., Abou-Khalil, R., & Roth, H. L. (2012). Toward a quantitative basis for assessment and diagnosis of apraxia of speech. *J Speech Lang Hear Res*, 55(5), S1502-1517. doi:10.1044/1092-4388(2012/11-0318)
- Haley, K. L., & Martin, G. (2011). Production variability and single word intelligibility in aphasia and apraxia of speech. *J Commun Disord*, 44(1), 103-115. doi:10.1016/j.jcomdis.2010.07.007
- Hall, J. W., & Grose, J. H. (1993). Monaural envelope correlation perception in listeners with normal hearing and cochlear impairment. *Journal of Speech, Language, and Hearing Research*, 36(6), 1306-1314.
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in quantitative methods for psychology*, 8(1), 23.

- Hartwigsen, G., Saur, D., Price, C. J., Baumgaertner, A., Ulmer, S., & Siebner, H. R. (2013). Increased facilitatory connectivity from the pre-SMA to the left dorsal premotor cortex during pseudoword repetition. *Journal of Cognitive Neuroscience*, 25(4), 580-594.
- Helm-Estabrooks, N. (1992). *Aphasia Diagnostic Profiles*. Austin, TX: Pro-Ed Inc.
- Hickok, G. (2012a). Computational neuroanatomy of speech production. *Nat Rev Neurosci*, 13(2), 135-145.
- Hickok, G. (2012b). The cortical organization of speech processing: feedback control and predictive coding the context of a dual-stream model. *J Commun Disord*, 45(6), 393-402. doi:10.1016/j.jcomdis.2012.06.004
- Hickok, G. (2014a). The architecture of speech production and the role of the phoneme in speech processing. *Lang Cogn Process*, 29(1), 2-20. doi:10.1080/01690965.2013.834370
- Hickok, G. (2014b). Toward an Integrated Psycholinguistic, Neurolinguistic, Sensorimotor Framework for Speech Production. *Lang Cogn Process*, 29(1), 52-59. doi:10.1080/01690965.2013.852907
- Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron*, 69(3), 407-422. doi:10.1016/j.neuron.2011.01.019
- Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends Cogn Sci*, 4(4), 131-138.
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1-2), 67-99. doi:10.1016/j.cognition.2003.10.011
- Hickok, G., Rogalsky, C., Chen, R., Herskovits, E. H., Townsley, S., & Hillis, A. E. (2014). Partially overlapping sensorimotor networks underlie speech praxis and verbal short-term memory: evidence from apraxia of speech following acute stroke. *Front Hum Neurosci*, 8, 649. doi:10.3389/fnhum.2014.00649
- Hillis, A. E. (1989). Efficacy and generalization of treatment for aphasic naming errors. *Arch Phys Med Rehabil*, 70(8), 632-636.
- Hillis, A. E., Work, M., Barker, P. B., Jacobs, M. A., Breese, E. L., & Maurer, K. (2004). Re-examining the brain regions crucial for orchestrating speech articulation. *Brain*, 127(Pt 7), 1479-1487. doi:10.1093/brain/awh172
- Hoit-Dalgaard, J., Murry, T., & Kopp, H. G. (1983). Voice onset time production and perception in apraxic subjects. *Brain and language*, 20(2), 329-339.
- Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience*, 5.
- Hyafil, A., Fontolan, L., Kabdebon, C., Gutkin, B., & Giraud, A.-L. (2015). Speech encoding by coupled cortical theta and gamma oscillations. *Elife*, 4, e06213.
- Ito, P. K. (1980). 7 Robustness of ANOVA and MANOVA test procedures. *Handbook of statistics*, 1, 199-236.
- Itoh, M., & Sasanuma, S. (1984). Articulatory Movements in Apraxia of Speech. In J. C. Rosenbek, M. R. McNeil, & A. E. Aronson (Eds.), *Apraxia of speech: Physiology, acoustics, linguistics, management* (pp. 135-165). San Diego, CA: College-Hill Press.

- Itoh, M., Sasanuma, S., Tatsumi, I. F., Murakami, S., Fukusako, Y., & Suzuki, T. (1982). Voice onset time characteristics in apraxia of speech. *Brain and Language*, 17(2), 193-210.
- Jacks, A. (2008). Bite block vowel production in apraxia of speech. *Journal of Speech, Language, and Hearing Research*, 51(4), 898-913.
- Johns, D. F., & Darley, F. L. (1970). Phonemic variability in apraxia of speech. *Journal of Speech and Hearing Research*, 13(3), 556-583.
- Jordan, L. C., & Hillis, A. E. (2006). Disorders of speech and language: aphasia, apraxia and dysarthria. *Curr Opin Neurol*, 19(6), 580-585.
doi:10.1097/WCO.0b013e3280109260
- Josephs, K. A., & Duffy, J. R. (2008). Apraxia of speech and nonfluent aphasia: a new clinical marker for corticobasal degeneration and progressive supranuclear palsy. *Current Opinion in Neurology*, 21(6), 688-692.
doi:10.1097/WCO.0b013e3283168ddd
- Josephs, K. A., Duffy, J. R., Strand, E. A., Machulda, M. M., Senjem, M. L., Gunter, J. L., . . . Whitwell, J. L. (2014). The evolution of primary progressive apraxia of speech. *Brain*, 137(Pt 10), 2783-2795. doi:10.1093/brain/awu223
- Josephs, K. A., Duffy, J. R., Strand, E. A., Machulda, M. M., Senjem, M. L., Lowe, V. J., . . . Whitwell, J. L. (2013). Syndromes dominated by apraxia of speech show distinct characteristics from agrammatic PPA. *Neurology*, 81(4), 337-345.
doi:10.1212/WNL.0b013e31829c5ed5
- Josephs, K. A., Duffy, J. R., Strand, E. A., Machulda, M. M., Senjem, M. L., Master, A. V., . . . Whitwell, J. L. (2012). Characterizing a neurodegenerative syndrome: primary progressive apraxia of speech. *Brain*, 135(Pt 5), 1522-1536.
doi:10.1093/brain/aws032
- Josephs, K. A., Duffy, J. R., Strand, E. A., Whitwell, J. L., Layton, K. F., Parisi, J. E., . . . Petersen, R. C. (2006). Clinicopathological and imaging correlates of progressive aphasia and apraxia of speech. *Brain*, 129(Pt 6), 1385-1398.
doi:10.1093/brain/awl078
- Kandel, E. R., Schwartz, J. H., & Jesell, T. M. (2000). *Principles of neural science* (4 ed.). New York, NY: McGraw Hill.
- Keating, P., & Shattuck-Hufnagel, S. (2002). A prosodic view of word form encoding for speech production. *UCLA working papers in phonetics*, 112-156.
- Kent, R. D., Kent, J. F., Duffy, J. R., Thomas, J. E., Weismer, G., & Stuntebeck, S. (2000). Ataxic dysarthria. *Journal of Speech, Language, and Hearing Research*, 43(5), 1275-1289.
- Kent, R. D., Kent, J. F., Weismer, G., & Duffy, J. R. (2000). What dysarthrias can tell us about the neural control of speech. *Journal of phonetics*, 28(3), 273-302.
- Kent, R. D., & McNeil, M. R. (1987). Relative timing of sentence repetition in apraxia of speech and conduction aphasia. In J. Ryalls (Ed.), *Phonetic approaches to speech production in aphasia and related disorders* (pp. 181-220). San Diego College Hill Press.
- Kent, R. D., & Rosenbek, J. C. (1983). Acoustic patterns of apraxia of speech. *Journal of Speech, Language, and Hearing Research*, 26(2), 231-249.
- Kertesz, A. (1982). *Western Aphasia Battery*. New York, NY: Grune & Stratton.
- Kertesz, A. (2007). *Western Aphasia Battery-Revised*. San Antonio, TX: Pearson.

- Kessinger, R. H., & Blumstein, S. E. (1998). Effects of speaking rate on voice-onset time and vowel production: Some implications for perception studies. *Journal of phonetics*, 26(2), 117-128.
- Kodumuri, N., Sebastian, R., Davis, C., Posner, J., Kim, E. H., Tippet, D. C., . . . Hillis, A. E. (2016). The association of insular stroke with lesion volume. *Neuroimage: Clinical*.
- Kotz, S. A., & Schwartz, M. (2010). Cortical speech processing unplugged: a timely subcortico-cortical framework. *Trends in Cognitive Sciences*, 14(9), 392-399.
- Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling* (Vol. 11): Sage.
- Lambon Ralph, M. A., Ehsan, S., Baker, G. A., & Rogers, T. T. (2012). Semantic memory is impaired in patients with unilateral anterior temporal lobe resection for temporal lobe epilepsy. *Brain*, 135(Pt 1), 242-258. doi:10.1093/brain/awr325
- Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, 14(1), 41-104.
- Levelt, W. J. M. (1989). Speaking: From intention to articulation. MA: The MIT Press, Cambridge.
- Levelt, W. J. M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, 42(1), 1-22.
- Levelt, W. J. M. (1999). Models of word production. *Trends Cogn Sci*, 3(6), 223-232.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(01), 1-38.
- Levelt, W. J. M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, 50(1), 239-269.
- Ling, L. E., Grabe, E., & Nolan, F. (2000). Quantitative Characterizations of Speech Rhythm: Syllable-Timing in Singapore English. *Language and speech*, 43(4), 377-401.
- Liss, J. M., LeGendre, S., & Lotto, A. J. (2010). Discriminating dysarthria type from envelope modulation spectra. *Journal of Speech, Language, and Hearing Research*, 53(5), 1246-1255.
- Liss, J. M., White, L., Mattys, S. L., Lansford, K., Lotto, A. J., Spitzer, S. M., & Caviness, J. N. (2009). Quantifying speech rhythm abnormalities in the dysarthrias. *Journal of Speech, Language, and Hearing Research*, 52(5), 1334-1352.
- Long, M. A., Katlowitz, K. A., Svirsky, M. A., Clary, R. C., Byun, T. M., Majaj, N., . . . Greenlee, J. D. (2016). Functional Segregation of Cortical Regions Underlying Speech Timing and Articulation. *Neuron*, 89(6), 1187-1193.
- Lowit, A. (2014). Quantification of rhythm problems in disordered speech: a re-evaluation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 369(1658), 20130404.
- Maas, E., Mailend, M.-L., & Guenther, F. H. (2015). Feedforward and Feedback Control in Apraxia of Speech: Effects of Noise Masking on Vowel Production. *Journal of Speech, Language, and Hearing Research*, 58(2), 185-200.
- Maas, E., Robin, D. A., Wright, D. L., & Ballard, K. J. (2008). Motor programming in apraxia of speech. *Brain Lang*, 106(2), 107-118. doi:10.1016/j.bandl.2008.03.004
- MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences*, 21(04), 499-511.

- Mailend, M.-L., & Maas, E. (2013). Speech motor programming in apraxia of speech: Evidence from a delayed picture-word interference task. *American Journal of Speech-Language Pathology*, 22(2), S380-S396.
- Martin, A. D. (1974). Some objections to the term apraxia of speech. *Journal of Speech and Hearing Disorders*, 39(1), 53-64.
- McNeil, M. R., Doyle, P. J., & Wambaugh, J. L. (2000). Apraxia of speech: A treatable disorder of planning and programming. In S. E. Nadeau, L. J. Gonzalenz Rothi, & B. Crosson (Eds.), *Aphasia and Language: Theory to Practice*. New York, NY: Guilford Press.
- McNeil, M. R., & Kent, R. D. (1990). Motoric characteristics of adult aphasic and apraxic speakers. *Advances in Psychology*, 70, 349-386.
- McNeil, M. R., Pratt, S. R., & Fossett, T. R. D. (2004). The Differential Diagnosis of Apraxia of Speech. In B. Maassen, R. Kent, H. Peters, P. van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 389-413). Oxford, UK: Oxford University Press.
- McNeil, M. R., Robin, D. A., & Schmidt, R. A. (1997). Apraxia of speech: Definition, differentiation, and treatment. In M. R. McNeil (Ed.), *Clinical Management of Sensorimotor Speech Disorders* (pp. 311-344). New York, NY: Thieme.
- Moser, D., Basilakos, A., Fillmore, P., & Fridriksson, J. (2016). Brain damage associated with apraxia of speech: evidence from case studies. *Neurocase*, 1-11. doi:10.1080/13554794.2016.1172645
- Moser, D., Fridriksson, J., Bonilha, L., Healy, E. W., Baylis, G., Baker, J. M., & Rorden, C. (2009). Neural recruitment for the production of native and novel speech sounds. *Neuroimage*, 46(2), 549-557.
- Mumby, K., Bowen, A., & Hesketh, A. (2007). Apraxia of speech: how reliable are speech and language therapists' diagnoses? *Clinical rehabilitation*, 21(8), 760-767.
- Munson, B. (2001). Phonological pattern frequency and speech production in adults and children. *Journal of Speech, Language, and Hearing Research*, 44(4), 778-792.
- Nachev, P., Coulthard, E., Jäger, H. R., Kennard, C., & Husain, M. (2008). Enantiomorphic normalization of focally lesioned brains. *Neuroimage*, 39(3), 1215-1226. doi:10.1016/j.neuroimage.2007.10.002
- Nadeau, S. (2000). *Phonology* (S. Nadeau, L. J. Gonzalenz Rothi, & B. Crosson Eds.). New York, NY: The Guilford Press.
- Nagao, M., Takeda, K., Komori, T., Isozaki, E., & Hirai, S. (1999). Apraxia of speech associated with an infarct in the precentral gyrus of the insula. *Neuroradiology*, 41(5), 356-357.
- NINDS. (2014). National Institute of Neurological Disorders and Stroke (NINDS) Aphasia Information Page. Retrieved from <http://www.ninds.nih.gov/disorders/aphasia/aphasia.htm>
- Niziolek, C. A., & Guenther, F. H. (2013). Vowel category boundaries enhance cortical and behavioral responses to speech feedback alterations. *Journal of Neuroscience*, 33(29), 12090-12098. doi:10.1523/JNEUROSCI.1008-13.2013
- Norton, A., Zipse, L., Marchina, S., & Schlaug, G. (2009). Melodic intonation therapy. *Ann N Y Acad Sci*, 1169(1), 431-436.

- Odell, K., McNeil, M. R., Rosenbek, J. C., & Hunter, L. (1991). Perceptual characteristics of vowel and prosody production in apraxic, aphasic, and dysarthric speakers. *Journal of Speech, Language, and Hearing Research*, 34(1), 67-80.
- Odell, K., McNeil, M. R., Rosenbek, J. C., & Hunter, L. (1991). Perceptual characteristics of vowel and prosody production in apraxic, aphasic, and dysarthric speakers. *J Speech Hear Res*, 34(1), 67-80.
- Ogar, J., Slama, H., Dronkers, N., Amici, S., & Gorno-Tempini, M. L. (2005). Apraxia of speech: an overview. *Neurocase*, 11(6), 427-432. doi:10.1080/13554790500263529
- Ogar, J., Willock, S., Baldo, J., Wilkins, D., Ludy, C., & Dronkers, N. (2006). Clinical and anatomical correlates of apraxia of speech. *Brain Lang*, 97(3), 343-350. doi:10.1016/j.bandl.2006.01.008
- Ojemann, G. A., & Whitaker, H. A. (1978). Language localization and variability. *Brain and language*, 6(2), 239-260.
- Okada, K., & Hickok, G. (2006). Identification of lexical-phonological networks in the superior temporal sulcus using functional magnetic resonance imaging. *Neuroreport*, 17(12), 1293-1296. doi:10.1097/01.wnr.0000233091.82536.b2
- Ordin, M., & Polyanskaya, L. (2015). Perception of speech rhythm in second language: the case of rhythmically similar L1 and L2. *Front Psychol*, 6.
- Pa, J., & Hickok, G. (2008). A parietal-temporal sensory-motor integration area for the human vocal tract: evidence from an fMRI study of skilled musicians. *Neuropsychologia*, 46(1), 362-368. doi:10.1016/j.neuropsychologia.2007.06.024
- Park, H., Rogalski, Y., Rodriguez, A. D., Zlatar, Z., Benjamin, M., Harnish, S., . . . Reilly, J. (2011). Perceptual cues used by listeners to discriminate fluent from nonfluent narrative discourse. *Aphasiology*, 25(9), 998-1015.
- Perkell, J. S. (2012). Movement goals and feedback and feedforward control mechanisms in speech production. *Journal of Neurolinguistics*, 25(5), 382-407.
- Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M., & Zandipour, M. (2004). The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *J Acoust Soc Am*, 116(4 Pt 1), 2338-2344.
- Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *J Acoust Soc Am*, 32(6), 693-703.
- Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech communication*, 41(1), 245-255.
- Pohar, M., Blas, M., & Turk, S. (2004). Comparison of logistic regression and linear discriminant analysis: a simulation study. *Metodoloski zvezki*, 1(1), 143.
- Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*, 62(2), 816-847.
- Racette, A., Bard, C., & Peretz, I. (2006). Making non-fluent aphasics speak: sing along! *Brain*, 129(10), 2571-2584.
- Richardson, J. D., Fillmore, P., Rorden, C., Lapointe, L. L., & Fridriksson, J. (2012). Re-establishing Broca's initial findings. *Brain Lang*, 123(2), 125-130. doi:10.1016/j.bandl.2012.08.007

- Roach, A., Schwartz, M. F., Martin, N., Grewal, R. S., & Brecher, A. (1996). The Philadelphia naming test: Scoring and rationale. *Clinical Aphasiology*, 24, 121-134.
- Robey, R. R. (1994). The efficacy of treatment for aphasic persons: a meta-analysis. *Brain Lang*, 47(4), 582-608. doi:10.1006/brln.1994.1060
- Robey, R. R. (1998). A meta-analysis of clinical outcomes in the treatment of aphasia. *J Speech Lang Hear Res*, 41(1), 172-187.
- Rogers, M. A. (1997). The vowel lengthening exaggeration effect in speakers with apraxia of speech: Compensation, artifact, or primary deficit? *Aphasiology*, 11(4-5), 433-445.
- Rorden, C., Bonilha, L., Fridriksson, J., Bender, B., & Karnath, H. O. (2012). Age-specific CT and MRI templates for spatial normalization. *Neuroimage*, 61(4), 957-965. doi:10.1016/j.neuroimage.2012.03.020
- Rorden, C., & Karnath, H. O. (2004). Using human brain lesions to infer function: a relic from a past era in the fMRI age? *Nat Rev Neurosci*, 5(10), 813-819. doi:10.1038/nrn1521
- Rosenbek, J. C., Kent, R. D., & Lapointe, L. L. (1984). Apraxia of speech: An overview and some perspectives. In J. C. Rosenbek, M. R. McNeil, & A. E. Aronson (Eds.), *Apraxia of speech: Physiology, acoustics, linguistics, management* (pp. 1-72). San Diego, CA: College-Hill Press.
- Rosenbek, J. C., McNeil, M. R., & Aronson, A. E. (1984). *Apraxia of speech : physiology, acoustics, linguistics, management*. San Diego, Calif.: College-Hill Press.
- Rosenbek, J. C., Wertz, R. T., & Darley, F. L. (1973). Oral sensation and perception in apraxia of speech and aphasia. *J Speech Hear Res*, 16(1), 22-36.
- Rubin, R. (2014). Stroke Care Projected to Add Up to More Than Previously Thought. *JAMA*, 312(21), 2198-2200.
- Schirmer, A. (2004). Timing speech: a review of lesion and neuroimaging findings. *Cognitive Brain Research*, 21(2), 269-287.
- Schlaug, G., Marchina, S., & Norton, A. (2008). From singing to speaking: why singing may lead to recovery of expressive language function in patients with Broca's aphasia. *Music perception: An interdisciplinary journal*, 25(4), 315-323.
- Schwartz, M. F., Faseyitan, O., Kim, J., & Coslett, H. B. (2012). The dorsal stream contribution to phonological retrieval in object naming. *Brain*, 135(Pt 12), 3799-3814. doi:10.1093/brain/aww300
- Seddoh, S. A., Robin, D. A., Sim, H. S., Hageman, C., Moon, J. B., & Folkins, J. W. (1996). Speech timing in apraxia of speech versus conduction aphasia. *J Speech Hear Res*, 39(3), 590-603.
- Shattuck-Hufnagel, S. (1979). *Speech errors as evidence for a serial-ordering mechanism in sentence production*. Paper presented at the Sentence processing: Psycholinguistic studies presented to Merrill Garrett.
- Shattuck - Hufnagel, S. (2015). Prosodic Frames in Speech Production. *The Handbook of Speech Production*, 419-444.
- Shewan, C. M., Leeper, H., & Booth, J. (1984). An analysis of voice onset time (VOT) in aphasic and normal subjects. *Apraxia of speech*, 197-220.

- Silbert, L. J., Honey, C. J., Simony, E., Poeppel, D., & Hasson, U. (2014). Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proceedings of the National Academy of Sciences*, 111(43), E4687-E4696.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., . . . Flitney, D. E. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, 23, S208-S219.
- Sörös, P., Sokoloff, L. G., Bose, A., McIntosh, A. R., Graham, S. J., & Stuss, D. T. (2006). Clustered functional MRI of overt speech production. *Neuroimage*, 32(1), 376-387.
- Spencer, K. A., & Slocumb, D. L. (2007). The neural basis of ataxic dysarthria. *The Cerebellum*, 6(1), 58-65.
- Sporns, O., Tononi, G., & Kötter, R. (2005). The human connectome: a structural description of the human brain. *PLoS Comput Biol*, 1(4), e42.
- SPSS. (2001). The SPSS TwoStep Cluster Component. Retrieved from http://www.spss.ch/upload/1122644952_The%20SPSS%20TwoStep%20Cluster%20Component.pdf
- Strand, E. A., Duffy, J. R., Clark, H. M., & Josephs, K. (2014). The apraxia of speech rating scale: a tool for diagnosis and description of apraxia of speech. *J Commun Disord*, 51, 43-50. doi:10.1016/j.jcomdis.2014.06.008
- Thomas, E. R., & Carter, P. M. (2006). Prosodic rhythm and African American English. *English World-Wide*, 27(3), 331-355.
- Tilsen, S., & Johnson, K. (2008). Low-frequency Fourier analysis of speech rhythm. *J Acoust Soc Am*, 124(2), EL34-EL39.
- Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Lang Cogn Process*, 26(7), 952-981. doi:10.1080/01690960903498424
- Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *Neuroimage*, 39(3), 1429-1443. doi:10.1016/j.neuroimage.2007.09.054
- Trupe, L. A., Varma, D. D., Gomez, Y., Race, D., Leigh, R., Hillis, A. E., & Gottesman, R. F. (2013). Chronic apraxia of speech and Broca's area. *Stroke*, 44(3), 740-744. doi:10.1161/STROKEAHA.112.678508
- Turk, A., & Shattuck-Hufnagel, S. (2013). What is speech rhythm? A commentary on Arvaniti and Rodriquez, Krivokapić, and Goswami and Leong. *Laboratory Phonology*, 4(1), 93-118.
- Turk, A., & Shattuck-Hufnagel, S. (2014). Timing in talking: what is it used for, and how is it controlled? *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 369(1658), 20130395.
- Ueno, T., & Lambon Ralph, M. A. (2013). The roles of the "ventral" semantic and "dorsal" pathways in conduite d'approche: a neuroanatomically-constrained computational modeling investigation. *Front Hum Neurosci*, 7, 422. doi:10.3389/fnhum.2013.00422
- Ueno, T., Saito, S., Rogers, T. T., & Lambon Ralph, M. A. (2011). Lichtheim 2: synthesizing aphasia and the neural basis of language in a neurocomputational

- model of the dual dorsal-ventral language pathways. *Neuron*, 72(2), 385-396. doi:10.1016/j.neuron.2011.09.013
- van der Merwe, A. (1997). A theoretical framework for the characterization of pathological speech sensorimotor control. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (pp. 1-25). New York, NT: Thieme.
- Varley, R. A., & Whiteside, S. P. (2001). What is the underlying impairment in acquired apraxia of speech. *Aphasiology*, 15(1), 39-49.
- Varley, R. A., Whiteside, S. P., & Luff, H. (1999). Dual-route speech encoding in normal and apraxic speakers: Some durational evidence. *Journal of Medical Speech and Language Pathology*, 7, 127-132.
- Vergis, M. K., Ballard, K. J., Duffy, J. R., McNeil, M. R., Scholl, D., & Layfield, C. (2014). An acoustic measure of lexical stress differentiates aphasia and aphasia plus apraxia of speech after stroke. *Aphasiology*, 28(5), 554-575.
- Wambaugh, J. L., Doyle, P. J., Kalinyak, M. M., & West, J. E. (1996). A critical review of acoustic analyses of aphasic and/or apraxic speech. *Clinical Aphasiology*, 24, 35-63.
- Wambaugh, J. L., Duffy, J. R., McNeil, M. R., Robin, D. A., & Rogers, M. (2006). Treatment guidelines for acquired apraxia of speech: Treatment descriptions and recommendations. *Journal of Medical Speech-Language Pathology*(142), xxxv-lxvii.
- Wambaugh, J. L., Kalinyak-Fliszar, M. M., West, J. E., & Doyle, P. J. (1998). Effects of treatment for sound errors in apraxia of speech and aphasia. *Journal of Speech, Language, and Hearing Research*, 41(4), 725-743.
- Wambaugh, J. L., Nessler, C., Cameron, R., & Mauszycki, S. C. (2012). Acquired apraxia of speech: the effects of repeated practice and rate/rhythm control treatments on sound production accuracy. *American Journal of Speech-Language Pathology*, 21(2), S5-27. doi:10.1044/1058-0360(2011/11-0102)
- Wambaugh, J. L., Nessler, C., Cameron, R., & Mauszycki, S. C. (2013). Treatment for acquired apraxia of speech: examination of treatment intensity and practice schedule. *American Journal of Speech-Language Pathology*, 22(1), 84-102. doi:10.1044/1058-0360(2012/12-0025)
- Wambaugh, J. L., Wright, S., Nessler, C., & Mauszycki, S. C. (2014). Combined Aphasia and Apraxia of Speech Treatment (CAAST): Effects of a Novel Therapy. *Journal of Speech, Language, and Hearing Research*, 57(6), 2191-2207.
- Wambaugh, J. L., Wright, S., Nessler, C., & Mauszycki, S. C. (2014). Combined Aphasia and Apraxia of Speech Treatment (CAAST): effects of a novel therapy. *J Speech Lang Hear Res*, 57(6), 2191-2207. doi:10.1044/2014_JSLHR-L-14-0004
- Wertz, R., LaPointe, L., & Rosenbek, J. (1984). Apraxia of speech: The disorders and its management. *New York: Grune and Stratton*, 318.
- White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, 35(4), 501-522.
- Whitwell, J. L., Duffy, J. R., Strand, E. A., Machulda, M. M., Senjem, M. L., Gunter, J. L., . . . Josephs, K. A. (2013). Neuroimaging comparison of primary progressive apraxia of speech and progressive supranuclear palsy. *Eur J Neurol*, 20(4), 629-637. doi:10.1111/ene.12004

- Whitwell, J. L., Duffy, J. R., Strand, E. A., Xia, R., Mandrekar, J., Machulda, M. M., . . . Josephs, K. A. (2013). Distinct regional anatomic and functional correlates of neurodegenerative apraxia of speech and aphasia: an MRI and FDG-PET study. *Brain Lang*, 125(3), 245-252. doi:10.1016/j.bandl.2013.02.005
- Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *Neuroimage*, 92, 381-397. doi:10.1016/j.neuroimage.2014.01.060
- Winkworth, A. L., Davis, P. J., Adams, R. D., & Ellis, E. (1995). Breathing patterns during spontaneous speech. *Journal of Speech, Language, and Hearing Research*, 38(1), 124-144.
- Wise, R., Greene, J., Büchel, C., & Scott, S. K. (1999). Brain regions involved in articulation. *The Lancet*, 353(9158), 1057-1061.
- Wu, O., Cloonan, L., Mocking, S. J., Bouts, M. J., Copen, W. A., Cougo-Pinto, P. T., . . . Rosand, J. (2015). Role of acute lesion topography in initial ischemic stroke severity and long-term functional outcomes. *Stroke*, 46(9), 2438-2444.
- Yorkston, K. M., & Beukelman, D. R. (1981). *Assessment of Intelligibility of Dysarthric Speech*. Tigard, OR: C.C. Publications.
- Ziegler, W. (2002). Psycholinguistic and motor theories of apraxia of speech. *Semin Speech Lang*, 23(4), 231-244. doi:10.1055/s-2002-35798
- Ziegler, W. (2005). A nonlinear model of word length effects in apraxia of speech. *Cogn Neuropsychol*, 22(5), 603-623. doi:10.1080/02643290442000211
- Ziegler, W. (2009). Modeling the architecture of phonetic plans: Evidence from apraxia of speech. *Language and Cognitive Processes*, 24, 631-661.
- Ziegler, W., Aichert, I., & Staiger, A. (2012). Apraxia of speech: concepts and controversies. *J Speech Lang Hear Res*, 55(5), S1485-1501. doi:10.1044/1092-4388(2012/12-0128)
- Ziegler, W., & von Cramon, D. (1985). Anticipatory coarticulation in a patient with apraxia of speech. *Brain Lang*, 26(1), 117-130.
- Ziegler, W., & von Cramon, D. (1986a). Disturbed coarticulation in apraxia of speech: acoustic evidence. *Brain Lang*, 29(1), 34-47.
- Ziegler, W., & von Cramon, D. (1986b). Timing deficits in apraxia of speech. *Eur Arch Psychiatry Neurol Sci*, 236(1), 44-49.
- Zlatin, M. A. (1974). Voicing contrast: Perceptual and productive voice onset time characteristics of adults. *J Acoust Soc Am*, 56(3), 981-994.

APPENDIX A

The Apraxia of Speech Rating Scale

Behaviors exclusive to AOS

- 1.1 Distorted sound substitutions
- 1.2 Distorted sound additions (not intrusive schwa)
- 1.3 Increased distortions/distorted sound substitutions with increased utterance length or increased syllable/word articulatory complexity
- 1.4 Increased sound distortions/distorted sound substitutions with increased speech rate
- 1.5 Inaccurate AMR's
- 1.6 Reduced words/breath group

Behaviors that can occur in AOS and Dysarthria

- 2.1 Syllable segmentation within words greater than one syllable
- 2.2 Syllable segmentation across words in phrases/sentences
- 2.3 Sound distortions
- 2.4 Slow overall rate
- 2.5 Lengthened vowel &/or consonant segments
- 2.6 Lengthened intersegment durations

Behaviors that can occur in AOS and aphasia

- 3.1 Deliberate, slowly sequenced, segmented, &/or distorted SMRs compared to AMRs
- 3.2 Audible/visual articulatory groping; initiation difficulty; false starts/restarts

Behaviors that can occur in AOS, aphasia, and dysarthria

- 4.1 Sound/syllable repetitions
- 4.2 Sound prolongations (beyond lengthened segments)

Strand, E. A., Duffy, J. R., Clark, H. M., & Josephs, K. (2014). The apraxia of speech rating scale: a tool for diagnosis and description of apraxia of speech. *J Commun Disord*, 51, 43-50. doi:10.1016/j.jcomdis.2014.06.008

APPENDIX B

Motor Speech Evaluation (Duffy, 2005) Supplemental Tasks for Assessing Motor Speech Abilities

- **Word repetition (3 times each)**
 1. cat
 2. catnip
 3. catapult
 4. catastrophe
 5. snowman
 6. artillery
 7. stethoscope
 8. rhinoceros
 9. volcano
 10. harmonica
 11. specific
 12. statistics
 13. aluminum
- **Max vowel prolongation**
- **Speech AMRs**
- **Speech SMRs (p-t-k)**
- **Word repetition** (repeat each word 3 times in succession; do not instruct to go fast or slow; provide a model with 3 repetitions of “boy”)
- **Sentence repetition (one repetition)**
 1. We saw several wild animals.
 2. My physician wrote out a prescription.
 3. The municipal judge sentenced the criminal.

Based on all speech data, address these questions...

1. Does the patient have dysarthria? Yes _____ No _____ Uncertain _____
2. If yes, what is the type? (circle those that apply) Flaccid Spastic Ataxic
Hypokinetic Hyperkinetic UUMN
3. Dysarthria Severity (0-4) _____
4. Is AOS present? Yes _____ No _____ Uncertain _____
5. AOS Severity (0-4) _____

APPENDIX C

Selected PNT items for word-level scoring

VOT Word List		rPVI Word List
<i>Voiced</i>	<i>Voiceless</i>	Ambulance (S-W)
Baby	Cake	Banana (W-S)
Balloon	Calendar	Binoculars (W-S)
Banana	Camera	Butterfly (S-W)
Basket	Can	Calendar (S-W)
Bat	Candle	Celery (S-W)
Beard	Carrot	Cheerleaders (S-W)
Belt	Cow	Dinosaur (S-W)
Bench	Key	Elephant (S-W)
Binoculars	King	Eskimo (S-W)
Bone	Pear	Helicopter (S-W)
Book	Peas	Microscope (S-W)
Boot	Pen	Octopus (S-W)
Bottle	Piano	Piano (W-S)
Bowl	Pie	Pineapple (S-W)
Bus	Pig	Pyramid (S-W)
Dice	Pillow	Stethoscope (S-W)
Dinosaur	Pineapple	Strawberries (S-W)
Dog	Pumpkin	Thermometer (W-S)
Door	Table	Typewriter (S-W)
Duck	Tent	Volcano (W-S)
Garage	Turkey	Waterfall (S-W)
Ghost	Cannon	
Goat	Comb	

Note: W-S indicates words with a weak-strong stress pattern; S-W indicates words with a strong-weak stress pattern

APPENDIX D

Lingual distortion errors used for narrow transcription (Cunningham et al., 2015)

1. Dentalised (tongue placed anteriorly at teeth)
2. Palatalised (tongue placement near hard palate)
3. Lateralised (tongue placement directs airflow laterally, rather than anteriorly)
4. Rhotacised (production of improper “r-coloring”)
5. Fronted (production at location anterior to target, e.g., /t/ produced anterior to palate)
6. Backed (production at location posterior to target, e.g., /t/ produced toward back of oral cavity)
7. Derhotacised (absence of “r-coloring”)
8. Frictionalised (airflow is not stopped adequately by tongue, e.g., /t/ produced with continuous airflow rather than plosive stop)

APPENDIX E

Regions chosen for exploratory connectome analysis

L IFGpo	R MTG	L mCereb Peduncle
R IFGpo	L MTG pole	R mCereb Peduncle
L IFGorb	R MTG pole	L PCT
R IFGorb	L ITG	R PCT
L IFGpt	R ITG	L iCereb Peduncle
R IFGpt	L Ins	R iCereb Peduncle
L PoCG	R Ins	L aCR
R PoCG	L Caud	R aCR
L PrCG	R Caud	L sCR
R PrCG	L Putamen	R sCR
L sPG	R Putamen	L pCR
R sPG	L GP	R pCR
L SMG	R GP	L pIns
R SMG	L Cerebellum	R pIns
L AG	R Cerebellum	L pSTG
R AG	L Cereb Peduncle	R pSTG
L STG	R Cereb Peduncle	L pMTG
R STG	L CST	R pMTG
L STG pole	R CST	L pITG
R STG pole	L sCereb Peduncle	R pITG
L MTG	R sCereb Peduncle	

APPENDIX F

Means (standard deviations) and significant group differences for each variable

	Stroke Control (S.C.)	Aphasia Only (A.O.)	AOS-Aphasia (AOS-A.)	Group Differences
Rhythm and Articulation				
nPVI-V	62 (3.44)	61.42 (5.60)	52.21 (7.11)	S.C. = A.O. S.C. > AOS-A A-O > AOS-A
VOT-SD_{voiced}	0.020 (0.007)	0.012 (0.01)	0.033 (0.03)	S.C. < AOS-A S.C. = A.O. A.O. = S.C.
VOT-SD_{voiceless}	0.026 (0.007)	0.03 (0.008)	0.033 (0.01)	S.C. = A.O. = AOS-A
Prop. Phonemic	0.002 (.003)	0.03 (.04)	0.11 (.09)	S.C. < A.O. < AOS-A.
Prop. Distortion	0.005 (.005)	0.02 (.03)	0.19 (.03)	S.C. = A.O. S.C. < AOS-A A-O < AOS-A
Rate				
Syllables/Second	1.15 (0.52)	0.06 (0.89)	-0.82 (0.46)	S.C. > A.O. S.C. > AOS-A. A.O. > AOS-A.
Mean WPM	1.12 (0.64)	0.034 (0.90)	-0.77 (0.49)	S.C. > A.O. > AOS-A.
Productivity	1.20 (0.66)	-0.11 (0.84)	-0.64 (0.63)	S.C. > A.O. S.C. > AOS-A. A.O. = AOS-A.
EMS Variables				
Peak Frequency	5.17 (1.34)	5.26 (1.93)	5.87 (3.47)	S.C. = A.O. = AOS-A.
Peak Amplitude	3.94 (0.89)	4.67 (1.23)	5.23 (1.18)	S.C. < A.O. S.C. < AOS-A. A.O. = AOS-A.
1 Hz Band	0.12 (0.01)	0.14 (0.04)	0.16 (0.02)	S.C. = A.O. S.C. < AOS-A. A.O. > AOS-A.
4 Hz Band	0.18 (0.02)	0.17 (0.02)	0.16 (0.02)	S.C. = A.O. S.C. > AOS-A. A.O. > AOS-A.
8 Hz Band	0.21 (0.01)	0.20 (0.02)	0.187 (0.017)	S.C. = A.O. S.C. < AOS-A. A.O. > AOS-A.
16 Hz Band	0.19 (0.01)	0.19 (0.02)	0.17 (0.02)	S.C. = A.O. S.C. > AOS-A. A.O. > AOS-A.
32 Hz Band	0.09 (0.01)	0.09 (0.01)	0.09 (0.02)	S.C. = A.O. = AOS-A.

APPENDIX G

Correlations between all measures

	nPVI- V	VOT- SD _{voiced}	Phon Err.	Dist. Err.	1 Hz	4 Hz	8 Hz	16 Hz	32 Hz	Peak Freq.
VOT- SD_{voiced}	-.41*									
<i>p</i> -value	0.002									
Phon Errors	-.48*	0.26								
<i>p</i> -value	<0.001	0.06								
Dist Errors	-.71*	.32*	.62*							
<i>p</i> -value	<0.001	0.02	<0.001							
1 Hz	-.35*	.31*	.49*	.53*						
<i>p</i> -value	0.007	0.02	<0.001	<0.001						
4 Hz	0.25	0.03	-.55*	-.34*	-.34*					
<i>p</i> -value	0.06	0.82	<0.001	0.01	0.01					
8 Hz	.50*	-0.26	-.41*	-.42*	-.70*	0.20				
<i>p</i> -value	<0.001	0.05	0.002	0.001	<0.001	0.13				
16 Hz	0.20	-0.20	-0.14	-.30*	-.53*	-0.11	.41*			
<i>p</i> -value	0.14	0.14	0.30	0.03	<0.001	0.42	0.002			
32 Hz	-0.08	-0.13	0.17	-0.10	-0.10	-.43*	0	0.62*		
<i>p</i> -value	0.56	0.34	0.21	0.48	0.45	0.001	1.0	<0.001		
Peak Freq.	-0.13	0.10	0.23	0.20	0.24	-.59*	0.02	.33*	.39*	
<i>p</i> -value	0.34	0.48	0.09	0.14	0.07	0	0.87	0.02	0.004	
Peak Amp.	0.11	0.05	-.45*	-.31*	-.51*	.52*	.32*	-0.13	-.40*	-.45*
<i>p</i> -value	0.40	0.70	<0.001	0.02	<0.001	<0.00	0.02	0.33	0.002	0.001

*Significant at $p < 0.05$, uncorrected

Correlations with Rate Measures

	Syll. Rate	WPM	Productivity
nPVI-V	.54*	.51*	.39*
<i>p</i>-value	<0.001	<0.001	0.002
VOT-SD_{voiced}	-.33*	-.36*	-.25
<i>p</i>-value	0.02	0.01	0.064
Dist Errors	-.71*	-.68*	-.70*
<i>p</i>-value	<0.001	<0.001	<0.001
Phon Errors	-.62*	-.58*	-.46*
<i>p</i>-value	<0.001	<0.001	<0.001
1 Hz	-.68*	-.66*	-.52*
<i>p</i>-value	<0.001	<0.001	<0.001
2 Hz	-0.32*	-0.34*	-0.32*
<i>p</i>-value	0.02	0.01	0.02
4 Hz	.58*	.54*	.60*
<i>p</i>-value	<0.001	<0.001	<0.001
8 Hz	.57*	.57*	.43*
<i>p</i>-value	<0.001	<0.001	0.001
16 Hz	.28*	0.26	0.15
<i>p</i>-value	0.04	0.05	0.26
32 Hz	-0.15	-0.15	-.28*
<i>p</i>-value	0.27	0.27	0.03
Peak Freq.	-0.14	-0.13	-0.21
<i>p</i>-value	0.31	0.35	0.12
Peak Amp.	-.53*	-.51*	-.64*
<i>p</i>-value	<0.001	<0.001	<0.001

*Significant at $p < 0.05$, uncorrected

Correlations with ASRS

	nPVI-V	VOT- SD_{voice}	Phon. Err.	Dist. Err.
ASRS1.1	-0.64*	0.41*	0.63*	0.78*
<i>p</i>-value	<0.001	0.002	<0.001	0<0.001
ASRS1.2	-0.50*	0.22	0.50*	0.62*
<i>p</i>-value	<0.001	0.119	<0.001	<0.001
ASRS1.3	-.68*	.37*	.57*	.80*
<i>p</i>-value	<0.001	0.006	<0.001	<0.001
ASRS1.4	-.67*	.37*	.55*	.75*
<i>p</i>-value	<0.001	0.006	<0.001	<0.001
ASRS1.5	-.57*	0.25	.55*	.66*
<i>p</i>-value	<0.001	0.072	<0.001	<0.001
ASRS1.6	-.58*	0.23	.52*	.65*
<i>p</i>-value	<0.001	0.1	<0.001	<0.001
ASRS2.1	-.65*	.33*	.58*	.69*
<i>p</i>-value	<0.001	0.016	<0.001	<0.001
ASRS2.2	-.73*	.38*	.64*	.77*
<i>p</i>-value	<0.001	0.005	<0.001	<0.001
ASRS2.3	-.73*	.43*	.66*	.81*
<i>p</i>-value	<0.001	0.002	<0.001	<0.001
ASRS2.4	-.74*	.36*	.56*	.75*
<i>p</i>-value	<0.001	0.009	<0.001	<0.001
ASRS2.5	-.62*	.36*	.58*	.67*
<i>p</i>-value	<0.001	0.008	<0.001	<0.001
ASRS2.6	-.63*	.28*	.58*	.72*
<i>p</i>-value	<0.001	0.04	<0.001	<0.001
ASRS3.1	-.59*	0.28	.59*	.57*
<i>p</i>-value	<0.001	0.053	<0.001	<0.001
ASRS3.2	-.52*	.31*	.73*	.73*
<i>p</i>-value	<0.001	0.024	<0.001	<0.001
ASRS4.1	-.43*	.31*	.56*	.49*
<i>p</i>-value	0.001	0.024	<0.001	<0.001
ASRS4.2	-0.02	-0.09	0.10	0.04
<i>p</i>-value	0.90	0.543	0.468	0.754

*Significant at $p<0.05$, uncorrected

Correlations with ASRS, Continued

	1 Hz	2 Hz	4 Hz	8Hz	16 Hz	32 Hz
ASRS1.1	.60*	0.19	-.43*	-.42*	-.31*	-0.14
p-value	<0.001	0.17	0.001	0.001	0.02	0.32
ASRS1.2	.44*	0.31*	-.32*	-0.23	-0.22	-0.18
p-value	0.001	0.03	0.017	0.10	0.11	0.18
ASRS1.3	.61*	0.23	-.41*	-.47*	-.36*	-0.18
p-value	<0.001	0.09	0.002	<0.001	0.01	0.20
ASRS1.4	.54*	0.25	-.37*	-.40*	-.316*	-0.07
p-value	<0.001	0.07	0.006	0.002	0.02	0.59
ASRS1.5	.54*	0.23	-.46*	-.36*	-.28*	-0.06
p-value	<0.001	0.10	0.001	0.01	0.05	0.65
ASRS1.6	.50*	0.19	-0.26	-.42*	-.32*	-.29*
p-value	<0.001	0.17	0.057	0.001	0.02	0.04
ASRS2.1	.48*	0.15	-0.26	-.43*	-.35*	-.28*
p-value	<0.001	0.27	0.06	0.001	0.01	0.04
ASRS2.2	.54*	0.19	-.29*	-.50*	-.35*	-0.22
p-value	<0.001	0.16	0.033	<0.001	0.01	0.11
ASRS2.3	.56*	0.22	-.35*	-.47*	-.39*	-0.14
p-value	<0.001	0.10	0.01	<0.001	0.004	0.31
ASRS2.4	.59*	0.20	-.33*	-.53*	-.38*	-0.13
p-value	<0.001	0.15	0.02	<0.001	0.01	0.36
ASRS2.5	.56*	0.20	-.33*	-.44*	-0.26	-0.15
p-value	<0.001	0.14	0.01	0.001	0.06	0.27
ASRS2.6	.52*	0.27	-.308*	-.37*	-.39*	-0.23
p-value	<0.001	0.047*	0.02	0.006	0.003	0.10
ASRS3.1	.49*	0.22	-.42*	-.46*	-.28*	0.03
p-value	<0.001	0.12	0.002	0.001	0.049	0.84
ASRS3.2	.76*	0.22	-.45*	-.60*	-.391*	0.05
p-value	<0.001	0.12	0.001	<0.001	0.003	0.71
ASRS4.1	.32*	0.23	-.47*	-.30*	-0.15	0.12
p-value	0.02	0.10	<0.001	0.03	0.27	0.40
ASRS4.2	-0.05	-0.21	-0.16	-0.02	0.20	0.23
p-value	0.75	0.13	0.26	0.91	0.15	0.11